

GIS をベースとした国勢調査のデータ公開の現状と課題

—日本と英国の比較を通して—

矢野桂司

1. はじめに

1980 年代後半に欧米で起こった GIS（地理情報システム）革命は、コンピュータで地図を効率よく描画するといった単なる新たな情報ツールの導入だけでなく、伝統的な地理学の研究方法や研究対象そのものを変化させてきた。そして、1990 年代中葉には、地理学を中心として、地図を扱う関連学問分野を巻き込んで、地理情報科学という新しい学問分野が形成された¹⁾。地理情報科学は、位置情報を含んだ地理空間情報である「データ」を、付加的な価値をもった「情報」に変え、それを分析することによって、新たな「知識」を創造することが期待されている²⁾。また、GIS 革命以降 4 半世紀が経過し、膨大かつ多様な地理情報が蓄積され、コンピュータの容量や処理速度が飛躍的に発達し、インターネットやモバイルなどの新たな ICT（情報コミュニケーション技術）の発展などもあり、GIS に基づいた、これまでにない新たな研究領域（例えば、IoT（Internet of Things）など）が切り拓かれている。

とりわけ、新たな地理空間情報としては、SNS、携帯電話、交通 IC カードなどのジオ・ビッグデータの出現に加え、これまでの官庁統計においても、GIS での利用を前提とした多くの地理空間情報が提供されるようになった。その背景には、日本では、2007 年の地理空間情報活用推進基本法の施行、国際的には、2013 年 6 月の G8 サミットでの、政府の持つデータのオープンな利活用に関するオープンデータ憲章での合意などがある³⁾。

例えば、日本の代表的な基幹統計のデータも 2010（平成 22）年国勢調査の結果から、そのほとんどが総務省統計局のホームページからダウンロードすることができるようになった。そこで、本研究では、日本の国勢調査データの公開状況を示し、英国における状況と比較しながら、日本の国勢調査のデータ公開に関する課題と今後の展開の可能性を論じることにする。

2. 日本における GIS ベースの国勢調査のデータ公開

1) 国勢調査データの公開

国勢調査は個人と世帯に関する基幹統計の 1 つである。その調査票は、これまで国勢調査員が受け持つおおむね 50 世帯の調査区単位に集められ、その住所情報に基づいて、全国、都道府県、市区町村といった空間単位で集計されて、様々な統計表が作成してきた。そして、1920（大正 9）年の第 1 回国勢調査以降、基本 5 年ごとに、『国勢調査報告』の冊子体として公開してきた。

市区町村よりも細かな空間単位としては、1970（昭和 45）年から地域メッシュ統計の 3 次メッシュ（基準地域メッシュ・1km メッシュ）によるデータが提供され、さらに、人口 20 万人以上の都市、およ

び人口が 20 万人未満でも県庁所在地である都市の市域内を細分化した「国勢統計区」が設定された。この国勢統計区は、長期間にわたって区画の変更を要しないよう考慮され、人口が概ね 1 万人程度（5,000 人から 2 万人未満の範囲）となるよう設定された。その背景には、「町村合併促進法」（1953 年）、「新市町村建設促進法」（1956 年）によるいわゆる昭和の大合併により、市町村数が激減したことが影響した⁴⁾。すなわち、合併による市町村域の拡大によって、市区町村よりも細かな空間単位での統計が求められたためである。しかし、メッシュ統計や後述の基本単位区の充実により、1980（昭和 55）年国勢調査を最後に集計されなくなった⁵⁾。

そして、1990（平成 2）年からは、恒久的な街区に相当する「基本単位区」が設定され（概ね 20-30 世帯）、1995（平成 7）年国勢調査から市区町村を町丁・字等別に細分化した「町丁・字等別集計」が提供されるようになった。もちろん、空間単位が増加することにより、統計表は大きくなり、もはや冊子体といった紙媒体での公表は不可能となり、GIS を用いなければ地図表示は困難となる。なお、冊子体の報告書の頒布は一般社団法人日本統計協会で、そして、電磁的記録媒体（CD-R 等）の頒布は公益財団法人統計情報研究開発センター（Sinfonica）で、それぞれ有償で行われている⁶⁾。

しかし、これらの国勢調査結果は、「e-Stat 政府統計の総合窓口（e-Stat）」において、ほとんどのものがホームページから公開されるようになった。2008（平成 20）年 4 月から、政府は、総務省を中心に、国民にとって便利で使いやすい統計データの提供や各府省等の統計情報システムの集約を図るために、「政府統計共同利用システム」をスタートさせ、その機能の 1 つとして、この「政府統計の総合窓口（e-Stat）」が運用されている⁷⁾。

その中で、国勢調査は、1920（大正 9）年の第 1 回調査から、1980（昭和 55）年までの国勢調査報告は PDF で、1985（昭和 60）年以降は CSV 形式、Excel 形式、DB 形式でダウンロードできるようになった。PDF 形式で提供されたものは、冊子体をスキャンしたもので、CSV 形式など機械判読可能な形式に変換する必要がある。e-Stat からは、CSV 形式以外にも Excel 形式のものと DB 形式のものが提供されている。DB 形式のものは、データ量（データのセル数）の上限が設けられているが、都道府県ごとに分かれ表を統合させたり、必要な項目だけを抽出したりすることが可能となる。しかし、多くの GIS ソフトの属性データの形式となる DBF に準じた形式（1 行目にフィールド名をもった形式）に加工する必要がある。なお、都道府県や市区町村の主題図を作成する場合は、GIS の地図データである都道府県ないし市区町村の境界データ（ポリゴン）とのマッチングのために、調査年月日の時点での、都道府県で 2 ヶタの、あるいは市区町村で 5 ヶタの全国地方公共団体コードを属性データに付与する必要がある。

これまで述べてきた様々な空間単位での国勢調査の GIS による地図化には、それらの空間単位に対応した境域の地図データが必要となる。市区町村の境域データとしては、国土数値情報ダウンロードサービスの「行政区域データ」がある⁸⁾。過去の境域に関しても、1920（大正 9）年をはじめ、1950（昭和 25）年から 2005（平成 17）年までは 5 年ごとに、そして 2006（平成 18）年以降は毎年の 1 月 1 日時点での市区町村境域が shape ファイルで提供されている。しかし、市区町村の合併・分区は 1 月 1 日とは限らず、国勢調査の 10 月 1 日の市区町村境域と一致しない可能性もある。この問題を解決するために、桐村ほかは年月日を指定することにより、当該の市区町村境域の GIS データを無償で生成する Web サービス「Municipality Map Maker for Web 市区町村区域の GIS データ生成ツール」を構築し、無償提供している⁹⁾。

国勢調査の市区町村より細かな集計単位としては、人口集中地区、地域メッシュ統計、町丁・字

等、基本単位区がある。市区町村の境域内で、人口密度の高い調査区（1990（平成2）年からは基本単位区）（原則として人口密度が1km²当たり4,000人以上）が隣接し、かつ、それらの地域の人口が5,000人以上を有する人口集中地区の地図データは、国土数値情報「人口集中地区データ」から1960（昭和35）年以降、5年おきに提供されている。

また、地域メッシュ統計に関しては、1973（昭和48）年7月12日行政管理庁告示第143号に基づく「標準地域メッシュ」がベースとなる。標準地域メッシュは、「基準地域メッシュ」、「分割地域メッシュ」及び「統合地域メッシュ」の3種類と定めている。基準地域メッシュは、第1次地域区画を基に区画され、第1次地域区画は、緯度を40分間隔、経度を1度間隔に区分した区画で、これを縦横に8等分した区画が第2次地域区画、更にこれを縦横に10等分した区画が基準地域メッシュ（第3次地域区画）である。

基準地域メッシュより細かな地域区分として2分の1地域メッシュがある。2分の1地域メッシュは基準地域メッシュを縦横に2等分したもので、緯度差は15秒、経度差は22.5秒で、1辺の長さは約500mである。そして、2005（平成17）年国勢調査からは東京都特別区と政令市に関して、2分の1地域メッシュをさらに縦横に2等分した、緯度差は7.5秒、経度差は11.25秒で、1辺の長さは約250mの4分の1地域メッシュでの集計が行われるようになった。また、2010（平成22）年調査からは、その範囲を県庁所在都市まで広げている。

こうした経緯度を基準に作成された地域メッシュの区画は、2001年施行の測量法の一部改定により、日本の測地基準系が、日本測地系から世界測地系へ移行することになった。その結果、2005（平成17）年国勢調査から世界測地系に基づく地域メッシュ統計が作成されるようになり、東京付近で北西へ約450mずれることとなり、経年的な比較が難しくなった¹⁰⁾。

その対応として、2005（平成17）年国勢調査に関しては、両測地系での地域メッシュ統計が作成され、過去の2000（平成12）年と1995（平成7）年の国勢調査についても、世界測地系での地域メッシュ統計が新たに作成され、Sinfonicaより頒布されている。なお、地域メッシュ統計のGISデータは、4隅の経緯度を計算して作成できるが、「e-Stat」を介して、基準地域メッシュ（世界測地系1kmメッシュ）と分割地域メッシュ（世界測地系500mメッシュ）の境域データを提供している。また、ESRIジャパン（株）からは、ArcGIS上で任意の地域の任意のサイズのメッシュを作成するプログラムが提供されている¹¹⁾。

そして、1995（平成7）年国勢調査からは町丁・字等の小地域統計が提供されるようになるが、その地図データに対応する「e-Stat」の境界データは、2000（平成12）年国勢調査からのものが、日本測地系と世界測地系それぞれの測地系に関して、緯度経度と平面直角座標系の座標系での、Shape形式とGML形式で提供されている。しかし、一度でダウンロードできる範囲は市区町村単位であり、1995（平成7）年国勢調査小地域統計の境域データは、Sinfonicaより都道府県単位で購入する必要がある。

なお、提供されている町丁・字等の境域データと、前述のe-Statの国勢調査の小地域統計の統計表を属性データとしてGISソフト上で結合させる場合、境域データに含まれる11桁のユニークなKEY_CODEが必要となる（2010（平成22）年国勢調査では、約22万町丁・字等がある）。例えば、京都府京都市北区等持院北町であれば、26101015018であるが、最初の5桁は市区町村コード（26京都府、101京都市北区）、そして、6桁の町丁字コード（015018、等持院北町）から構成されている。しかし、e-Statで提供される統計表では、市区町村コードと町丁字コードが示されるものの、これらを文字

列として結合された 11 桁の町丁・字等の境域データの KEY_CODE に対応するものもなく、利用者が当該のコードをさらに発生させなければならない。

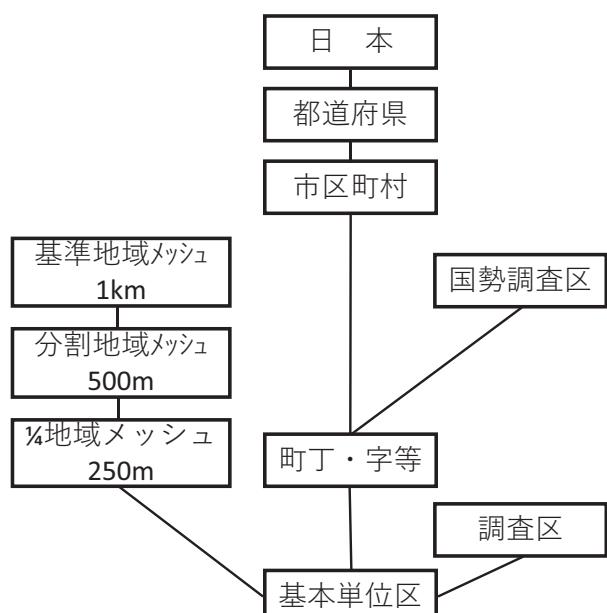
また、1990（平成 2）年国勢調査から設定された基本単位区の境域データは、「e-Stat」からは提供されておらず、Sinfonica から 1995（平成 7）年国勢調査以降の代表点のポイントデータが有償頒布されている。基本単位区や調査区の正確な境域は、統計局や各自治体において閲覧可能であるが、地図データとしての提供は行われていない。なお、昭和 25 年国勢調査から整備された調査区は、1975（昭和 50）年国勢調査から正確な地図が作成されているものの¹²⁾、調査区集計の統計表が提供されるのは 1980（昭和 55）年・1985（昭和 60）年国勢調査の 2 回のみである。これらの地図データは提供されていないが、調査区地図は統計局で閲覧することができる。

なお、基本単位区コードは、9 桁からなるもので最初の 6 桁は町丁字コードに対応している。例えば、先ほどの京都府京都市北区等持院北町は、2010（平成 22）年国勢調査時点では、4 つの基本単位区から構成されており、それぞれのコードは、015018010、015018020、015018030、015018040 となっている。したがって、2010（平成 22）年国勢調査で、全国に 1,885,188 ある基本単位区は、5 桁の市区町村コードと 9 桁の基本単位区コードを合わせた 14 桁で基本的に識別されることになる。

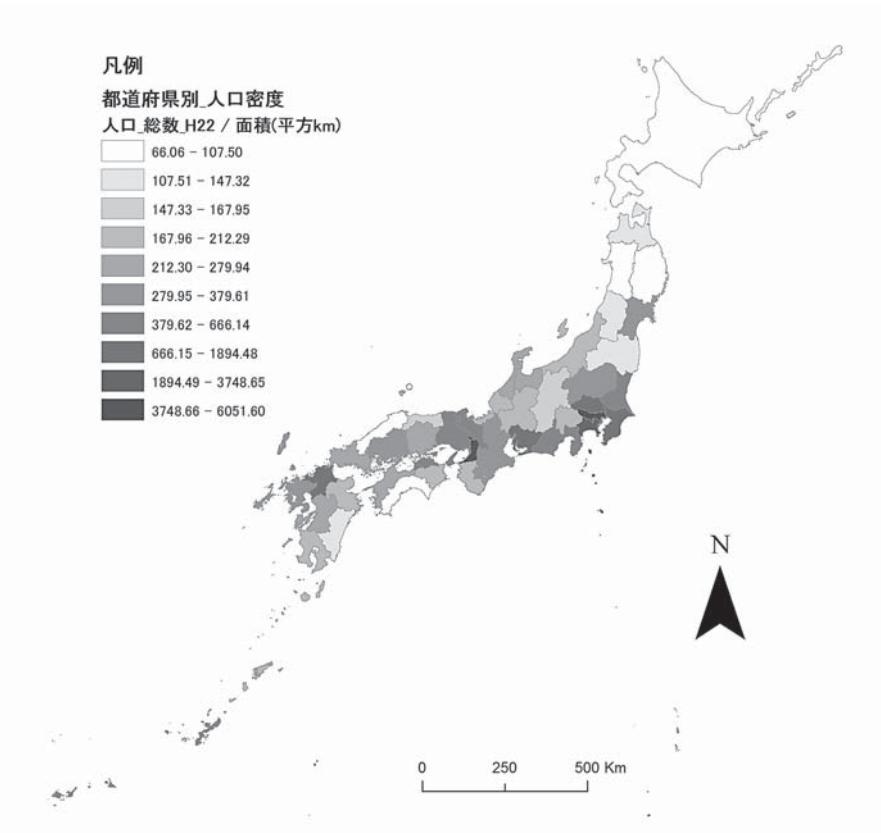
これまでの日本の国勢調査の統計表に対応した空間単位は第 1 図のようになる。また、これらの空間単位を、2010（平成 22）年国勢調査の統計表に対応させて地図化すると、第 2 図のようになる。ここでは、47 都道府県と 1907 市区町村の人口密度の階級区分図を描いている（第 2-a、b 図）。また、市区町村よりも細かな空間単位として、京都市を事例に、11 区（第 3-a 図）町丁・字等（第 3-b 図）、1km メッシュ（第 3-c 図）、500m メッシュ（第 3-d 図）、250m メッシュ（第 3-e 図）、基本単位区の代表点（第 3-e 図）での人口の階級区分図をそれぞれ地図化した。そして、最後に、京都市北区等持院北町を含むそれらの大きさを示したものが第 4 図である。このような非常に細かな空間単位で、全国すべての地域で国勢調査の変数が公開されていることは、十分に知られていない。また、これらの空間単位で集計される統計表は、それぞれの空間単位で異なるとともに、調査年次によっても公開される統計表と表章される変数は異なる点に注意する必要がある。さらに、調査年次によって、空間単位の境域も異なっており、経年的な変化を分析する場合はそれらの変化にも注意する必要がある。

2) 国勢調査の WebGIS

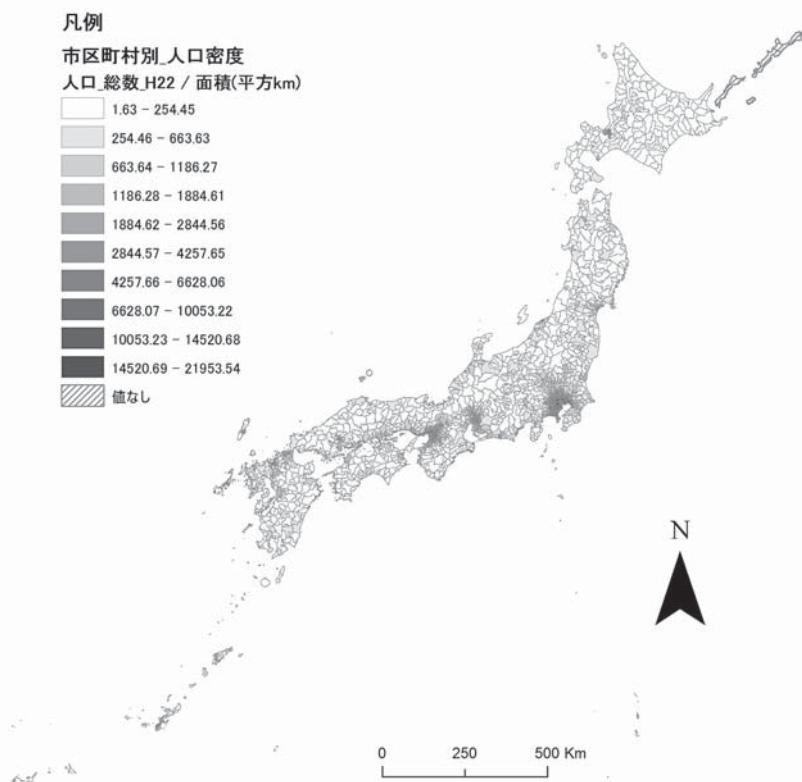
総務省時計局は、e-Stat の中で、WebGIS を「地図で見る統計（統計 GIS）」として提供している。そこでは、国勢調査以外の経済センサスなどの官庁統計の GIS 化も可能であるが、国勢調査に限定すれば、2010（平成 22）年、2005（平成 17）年、2000（平成 12）年国勢調査の都道府県集計、市区町村集計、小地域集計に関する特定の変数に関して、Web 上で主題図を作成することができ、階級区分図以外に、可変シンボルや棒グラフでの地図も作成する



第 1 図 日本国勢調査の集計空間単位の構成

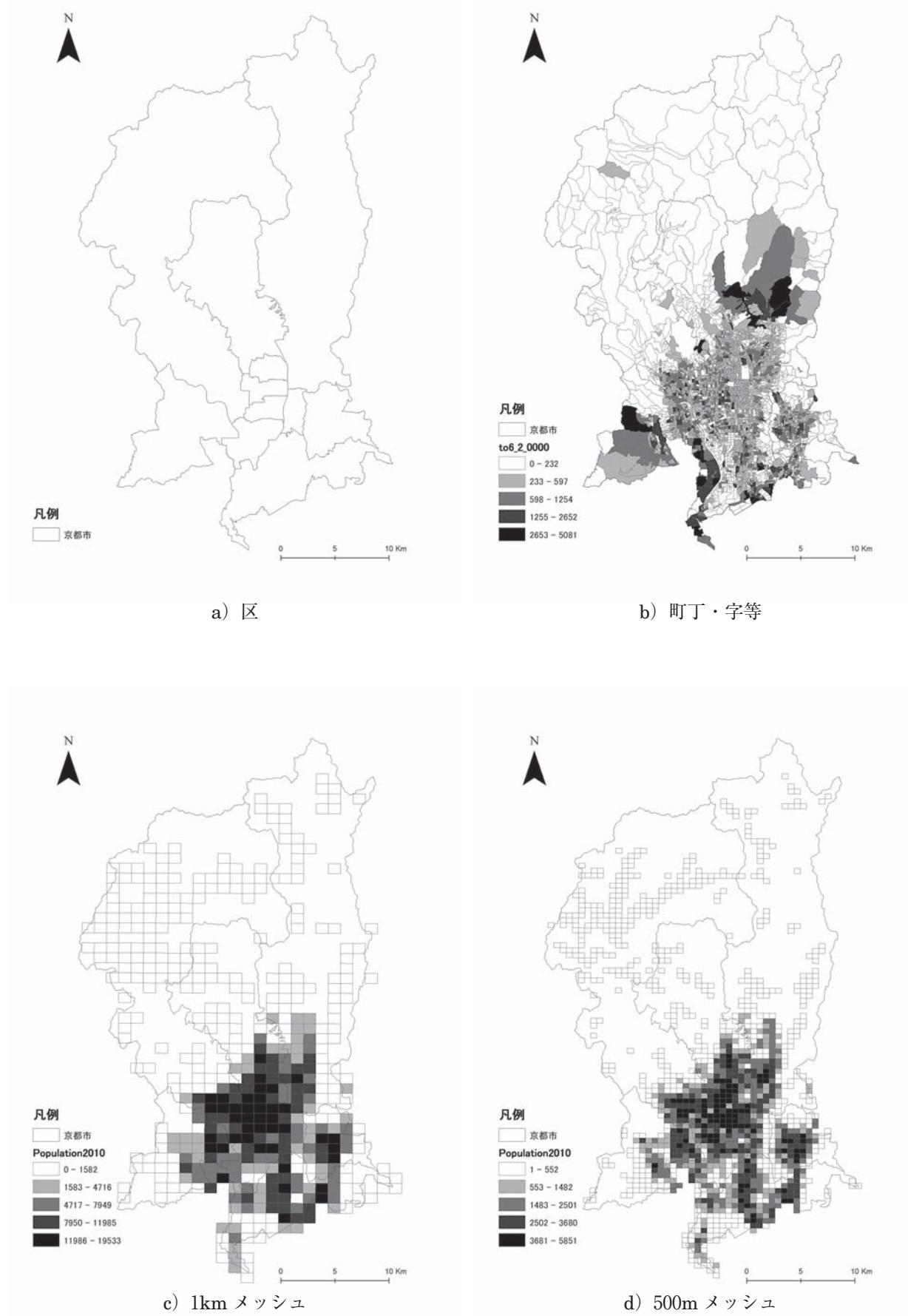


a) 都道府県

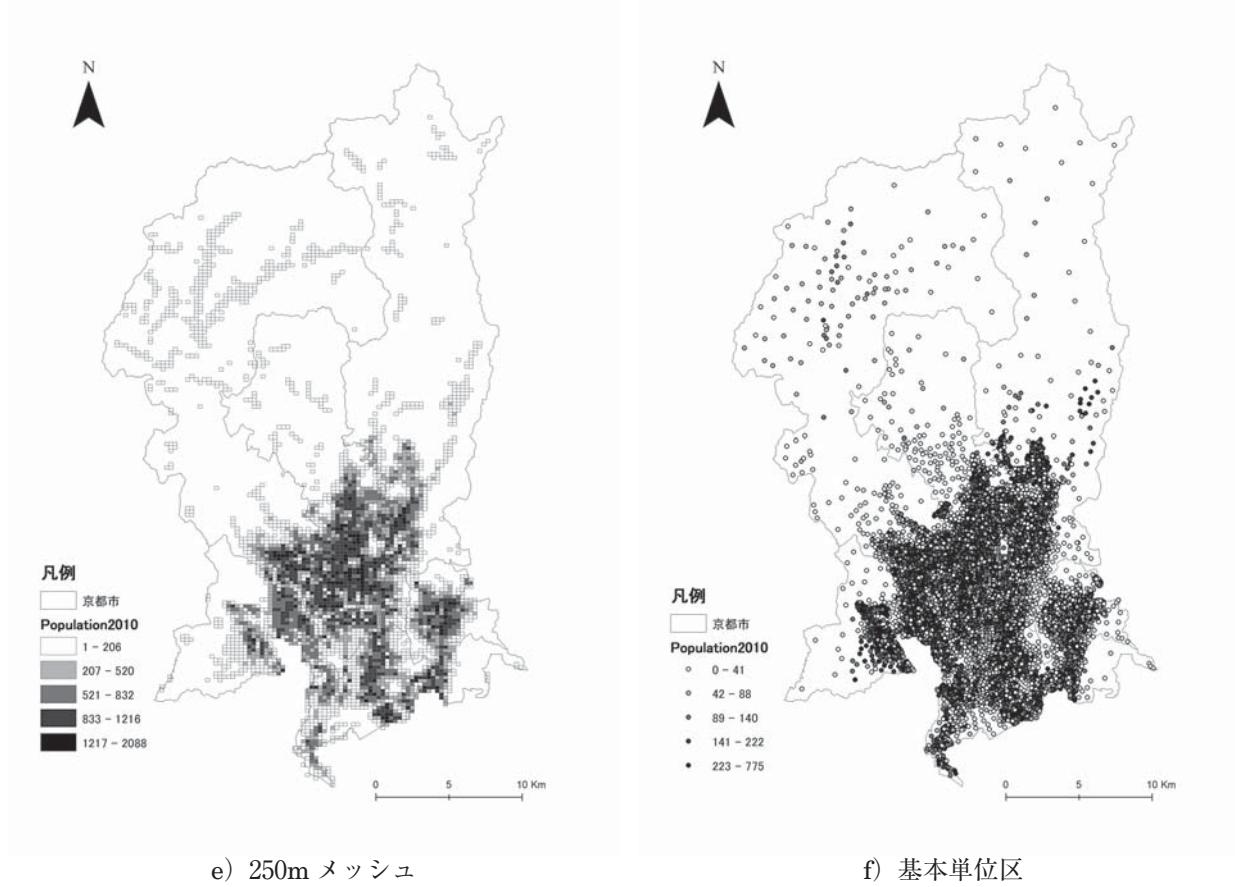


b) 市区町村

第2図 人口密度の都道府県・市区町村の階級区分図（2010（平成22）年国勢調査）



第3図 京都市の人口の階級区分図（2010（平成22）年国勢調査）



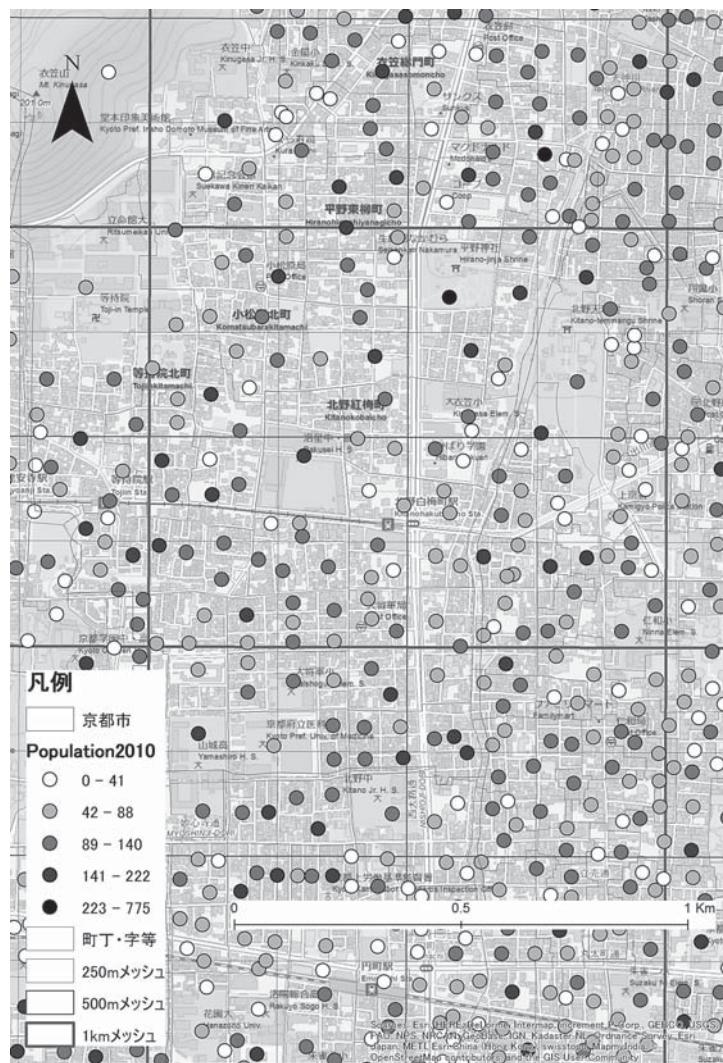
第3図 (続き)

ことができる。また、地域メッシュ統計は、平成7(1995)年・平成12(2000)年国勢調査に関しては、世界測地系の1kmメッシュと500mメッシュ、2005(平成17)年・2010(平成22)年国勢調査に関しては、それらに加えて、世界測地系の250mメッシュ(東京都特別区部、政令指定都市及び県庁所在地を含む第2次地区画の地域)も地図化できる。ただし、変数は男女別人口と世帯数のみと限定的である。

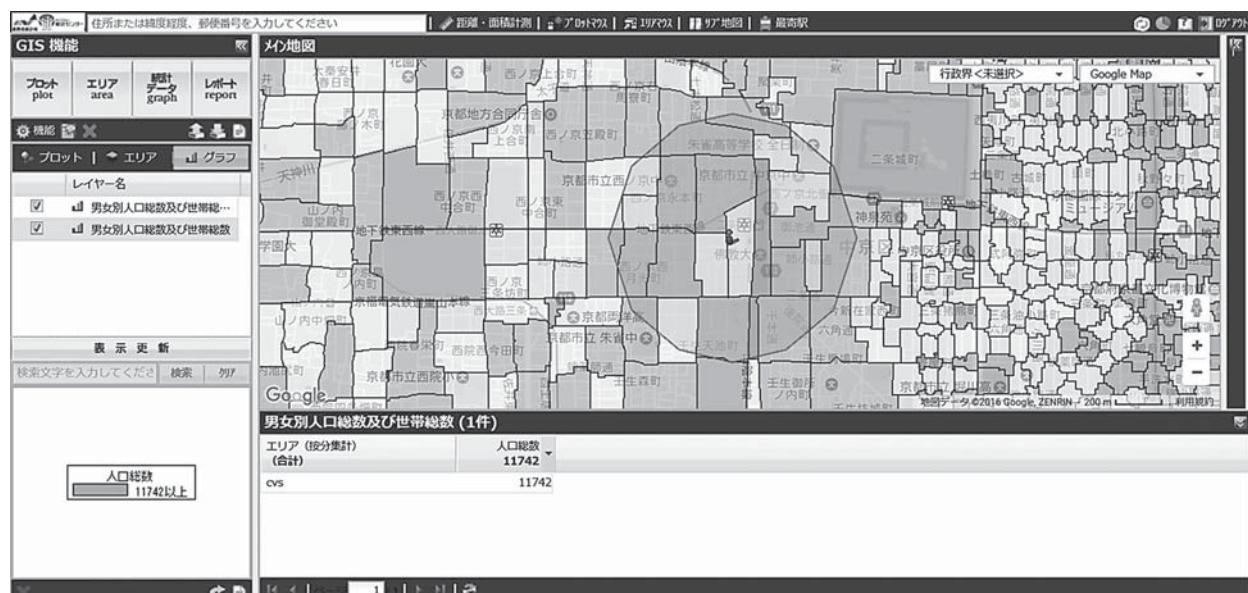
このシステムは、GISソフトを導入することなしに、主題図の階級区分数や階級の区切りや彩色を自由に選択できるなど、センサス・マッピングのシステム提供としては画期的である。しかし、選択できる変数が限定であることや、総人口による比率や単位面積当たりの密度などに加工する機能がないことなど、発展性や応用性に欠けているといえる。

さらに、これとは別に総務省統計局と統計センターは、「統計におけるオープンデータの高度化」の一環として、①利用者の保有するデータを取り込んで分析する機能、②任意に指定したエリアにおける統計算出機能、③地域分析レポート作成機能、を持ったWebGISである「地図による小地域分析(jSTAT MAP)」を提供している。このシステムは、Google Mapsと地理院地図を背景地図として、主題図を描くことができる。さらに、ジオコーディング機能や点やポリゴンなどの新たなフューチャーを追加したり、バッファリング機能も有している。これらの機能を用いれば、一般的なGISソフトで行うエリアマーケティングの分析も可能である。

例えば、Google Maps上のあるコンビニ上にポイントを発生させ、半径500mの円バッファを生成する。そして、町丁・字等別の総人口の主題図を描き、500m円バッファに含まれる総人口を集計



第4図 市区町村よりも細かな国勢調査の
空間単位の比較（2010（平成22）年国勢調査）
(京都市北区等持院北町周辺)



第5図 「地図による小地域分析 (jSTAT MAP)」のWeb画面

し、11,742 の商圈人口を推定することができる（第5図）。その集計方法も場合も、500m 円バッファにかかる町丁・字等の変数を面積按分する場合としない場合で集計することができる。さらに、レポート機能を用いると、当該の 500m 円バッファに重なる町丁・字等の 1) 基本分析、2) 周辺地図、3) かかる小地域、4) 年齢別人口、5) 世帯数、6) 経済センサス、7) 人口・世帯数増減、8) マップキャプチャ、を Excel ファイルとして出力することもできる。

3) 国勢調査のオープンデータ化の評価と課題

前述のように 2010（平成 22）年国勢調査の公表は、本格的な官庁統計のオープンデータを実践したものとして高く評価される。さらに、過去の国勢調査に関しても積極的に電子化を行っており、経年的な比較も可能としている。そして、「地図で見る統計（統計 GIS）」や「地図による小地域分析(jSTAT MAP)」などの WebGIS が整備され、多くの人が国勢調査の統計表を表やグラフで表示するだけでなく、容易にセンサス・マッピングの主題図として地図化することができるようになった点は大きい。

しかし、この WebGIS では、e-Stat で統計表として公開されている全ての変数を地図化できるわけではない。その場合は、当該の統計表と境域の地図データをダウンロードして、GIS ソフトによって地図化を行わなければならないが、その加工プロセスが必ずしも簡単ではない。

1995（平成 7）年国勢調査からはじまる小地域統計のオープンデータ化に関しても、2010（平成 22）年国勢調査の統計表（第1表から第20表）に合わせて公開された。小地域統計の統計表は、各調査年次で異なり、表章される変数も大きく異なる。大規模調査と簡易調査の違いを念頭に置いたとしても、変数の細分化が異なる。例えば、町丁・字等集計に関して、表章変数が最も多かった平成 12 年国勢調査においては、基本単位区集計で第1表から第3表の計 84 変数が、町丁・字等集計で第4表から第21表の計 5,625 変数が表章されている。しかし、平成 22 年国勢調査では、基本単位区集計で第1表の計 4 変数（総人口、男、女、世帯数）、町丁・字等集計で第2表から第20表の計 669 変数と、表章変数が大幅に減少している。小地域統計の場合、基本単位区で 189 万（約 20-30 世帯）、町丁・字等で約 22 万とデータ量が膨大となる。さらに、母数が少なくなり、個人情報保護の観点による秘匿データが多くなり分析に適さなくなってしまうという問題もある。

一方で、男女別年齢階級別の細かな分析が必要になることが多い。2007 年の新統計法により、統計データの利用促進として、統計の研究や教育など公益に資するために使用される場合に限り、二次的に利用することが可能となり、特に、学術研究目的、大学などの高等教育目的又は国際社会における我が国の利益の増進及び国際経済社会の健全な発展などのために、オーダーメードで集計された統計の提供を受けたり、匿名データ（調査票情報を個人や企業が特定できない形に加工したもの）の提供を受けて統計の作成に用いたりする委託による統計の作成、匿名データの作成・提供が可能となった。このように個人情報の保護とオーダーメード集計の提供といった観点を勘案しながら、適切なオープンデータの推進が進められる必要がある。

3. 英国の国勢調査のデータ公開

英国は、その正式名称がグレートブリテン及び北アイルランド連合王国であるように、英格ラ

ンド、ウェールズ、スコットランド、北アイルランドの4つの国 Country から構成されている。これら4つの異なる国の統一の国勢調査（人口センサス Population Census）は1801年に始まり、その後、10年おきに行われ、最新のものは2011年3月27日のものである。イングランドとウェールズはOffice for National Statistics (ONS)、スコットランドはNational Records of Scotland (NRS)、北アイルランドはthe Northern Ireland Statistics and Research Agency (NISRA)が、現在、それぞれの管轄における国勢調査を実施している¹³⁾。英国での国勢調査のデジタル化は、1971年調査から本格化し、主に英國の地理学者の研究グループが、センサス地理学 (Census Geography) を提唱して、大学と統計局が連携する形で、国勢調査のデジタルデータが構築され、提供されてきた¹⁴⁾。

1) 英国の国勢調査のWeb公開

これまで、1971、1981、1991、2001年国勢調査のデータは、Casweb (the web interface to Census Area Statistics and boundary data) から、利用登録を行った研究者に対して提供されてきた。Caswebは、1969年に社会科学研究会議SSRC (現在の経済・社会研究会議 (ESRC)) の基金によって設立されたUKDAの助成によって、Essex大学を中心に作成されたものである¹⁵⁾。このシステムは、ESRCの基金を受けて、2012年10月1日からUK Data Service¹⁶⁾に統合され、2011年センサスとともに、全ての集計データが一元的に提供されるようになった。

2001年以前の英国の国勢調査を提供するCaswebからは、第1表のものが提供されているが、4つの国Countryによって、集計される空間単位が異なる点にも注意が必要である。1991年センサスから集計表は、これまでのSAS (Small Area Statistics) だけでなく、SASとLBS (Local Base Statistics)の2つの集計表が作成されている。SASはLBSの一部で、同じ集計表でもカテゴリーが異なる。これらの違いは、過去の比較を可能とするSASの集計表と、コンピュータの発達による大容量データの処理能力により、多くの集計表が作成されたことによる¹⁷⁾。

そして、UK Data Serviceは、本格的なオープンデータの時代を迎えた2011年国勢調査に関しては、同じWebサイトからInFuseという国勢調査の無料提供サービスを行っている。InFuseでは、Kaswebと異なって、対象とする変数を住宅タイプ、年齢などの100を超えるTopics (2011年で166項目、2001年で96項目) から選択でき、他のGISソフトで利用可能なように、属性データと境域データとを合わせてダウンロードすることができる。2011年国勢調査に関しては838の、2001年国勢調査に関しては515のTopicsの組み合わせから集計表を選択することができ、集計する空間単位に関しても様々な空間スケールで選ぶことができる(第2表)。2011年国勢調査の空間単位は、第6図のようである。

2011年国勢調査の空間単位を地図化すると第7図のようになる。4つの国の行政区はそれぞれ異なるが、イングランドであれば、9つのRegion、55のCountyに分かれる。4つの国に共通する、統計区としてはDistrict (404)があり、ロンドン市は、32のロンドン自治区BoroughとCityに分かれる。それよりも細かな空間単位は、選挙区に対応するWard and Electoral Division (10,672)、さらにMiddle Super Output Areas and Intermediate Zones (8,480)、Lower Super Output Areas and Data Zones (42,619)、Output Areas and Small Areas (OA) (232,96)と階層化されている。OAの平均世帯数は約113.8である。

これまで述べてきた1971年以降の国勢調査の空間単位に対応した境域データも全てWeb上に用意されている。具体的には、2001年からのOpen Government Licence (OGLby)に則り、OGLに

第1表 英国の2001年、1991年、1981年、1971年国勢調査の統計表と境域データ

- 1) 2001 Aggregate Statistics Datasets (with digital boundary data)
- a) 統計表
- Key Statistics (KS001-KS024) 【提供空間単位、以下同様、CAS Ward、OA】
 Standard Tables (ST001-ST161) 【ST Ward】
 Standard Tables Theme Tables (TT001-TT053) 【ST Ward】
 Census Area Statistics (CS001-CS126) 【CAS Ward、OA】
 Census Area Statistics Theme Tables (CT001-CT011) 【CAS Ward、OA】
 Census Area Statistics Univariate Tables (UV001-UV093) 【CAS Ward、OA】
 Armed Forces Tables (AF001-AF004)
- b) 国別の提供空間単位
- England 【Country、GOR、County、Unitary Authority、District、ST Ward、CAS Ward、OA】
 Wales 【Country、Unitary Authority、ST Electoral Div、CAS Electoral Div、OA】
 Scotland 【Country、Council Area、ST Postcode Sector、CAS Postcode Sector、OA】
 Northern Ireland 【Country、District Council、Ward、OA】
- 2) 2001 Aggregate Statistics Datasets for Super Output Area (SOA) Geography
- a) 統計表
- Key Statistics (KS001-KS023) 【MSOA、LSOA、OA】
 Census Area Statistics Univariate Tables (UV001-UV085) 【MSOA、LSOA、OA】
- b) 国別の提供空間単位
- England 【MSOA、LSOA、OA】
 Wales 【MSOA、LSOA、OA】
 Scotland 【なし】
 Northern Ireland 【SOA、OA】
- 3) 1991 Great Britain SAS and LBS (with boundary data)
- a) 統計表
- Great Britain SAS (SAS01-99)
 Great Britain LBS (LBP01-99)
- b) 国別の提供空間単位
- England 【Country、District、Ward、ED】
 Wales 【Country、District、Ward、ED】
 Scotland 【Region、District、Postcode Sector、OA】
- 4) 1991 Northern Ireland SAS
- a) 統計表
- Northern Ireland SAS (SAS01-95)
- b) 提供空間単位
- Northern Ireland 【Northern Ireland、District Council、Ward、ED】
- 5) 1981 Great Britain SAS
- a) 統計表
- Great Britain SAS (SAS01-53)
- b) 国別の提供空間単位
- England 【Country、District、Ward、ED】
 Wales 【Country、District、Ward、ED】
 Scotland 【Region、District、Postcode Sector、OA】
- 6) 1971 Great Britain SAS
- a) 統計表
- Great Britain SAS (SAS01-28)
- b) 国別の提供空間単位
- England 【Country、Country、District、Ward、ED】
 Wales 【Country、District、Ward、ED】
 Scotland 【Region、District、Postcode Sector、OA】

第 2 表 英国の 2011 年国勢調査の統計表と境域データ

a) 統計表

2011 年国勢調査 (838 トピックス) に関しては、以下を参照。

<http://infuse2011.ukdataservice.ac.uk/InFuseWiz.aspx?cookie=openaccess>

2001 年国勢調査 (515 トピックス) に関しては。以下を参照。

<http://infuse.ukdataservice.ac.uk/infusewiz.aspx?cookie=openaccess>

b) 集計空間単位

Countries and Groupings (7 areas)

Regions (9 areas)

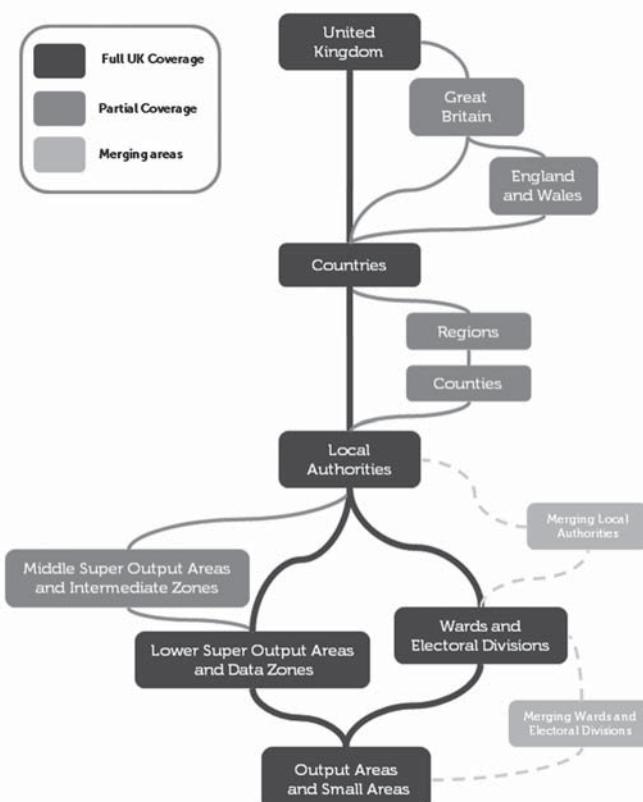
Local Authorities (404 areas)

Wards and Electoral Divisions (9481 areas)

Middle Super Output Areas and Intermediate Zones (8480 areas)

Lower Super Output Areas and Data Zones (42619 areas)

Output Areas and Small Areas (232296 areas)

第 6 図 英国の 2011 年国勢調査の集計空間単位の構成
(InFuse ホームページより転載)

準拠した境域データが、UK Data Service の Census Geography (<https://borders.ukdataservice.ac.uk/>) から無償で提供されている。

Census Geography のサイトは、いくつかのパートから構成されているが、その中の「Easy Download」 (https://census.edina.ac.uk/easy_download.html) からは、国勢調査のほとんど全ての境域データを簡単にダウンロードすることができる。境域データのデータ形式としては、SHAPE、KML、MAPINFO、CSV から選ぶことができる。さらに、「Boundary Data Selection」 (https://census.edina.ac.uk/boundary_data_selection.html)

ac.uk/bds.html) のサイトでは、英国政府が行う統計データの境域データを選べるうえに、Map View という WebGIS で、事前にその境域を確認することもできる。さらに、国勢調査の統計表の CSV ファイルを準備すれば、様々な階級区分図を描く WebGIS「Thematic Mapper」(<https://borders.ukdataservice.ac.uk/thematic/map>) も用意されている。

4. 日本と英国の国勢調査のデータ公開方法の違い

英国の国勢調査の統計表や境域データの公開と、日本の国勢調査の公開が大きく異なる点は、その公開主体の違いにある。日本の場合、総務省統計局が中心となってデータの



a) Countries (4)

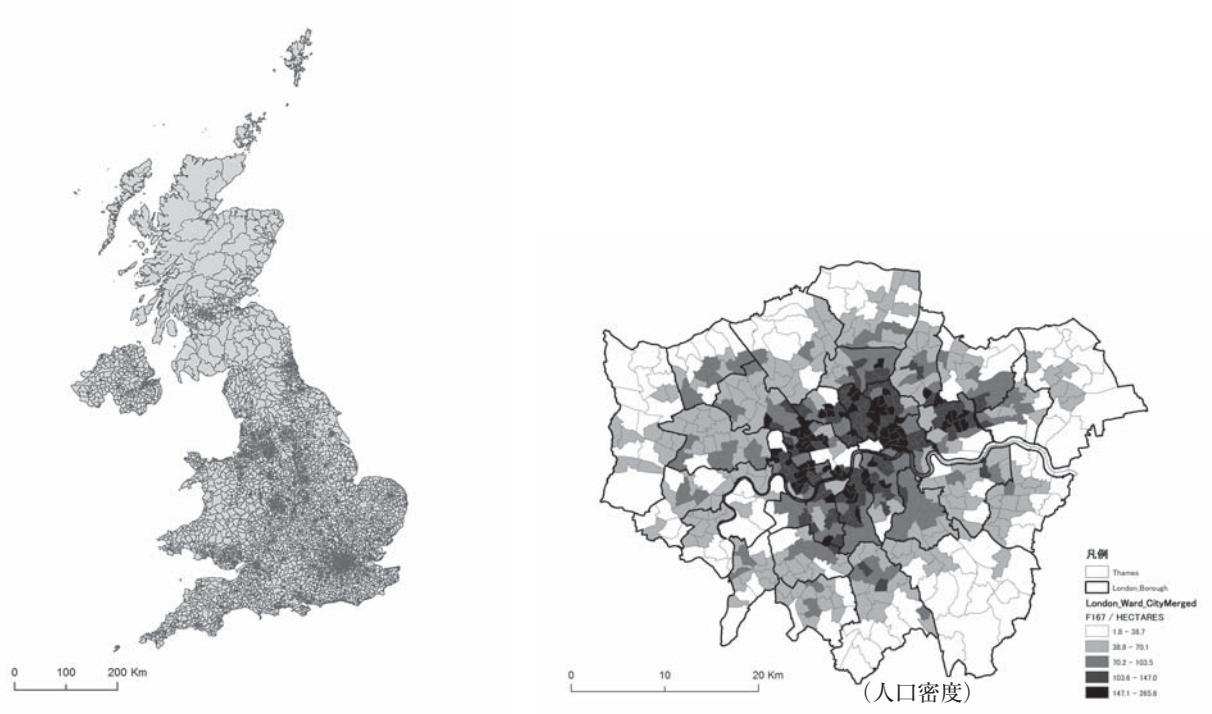
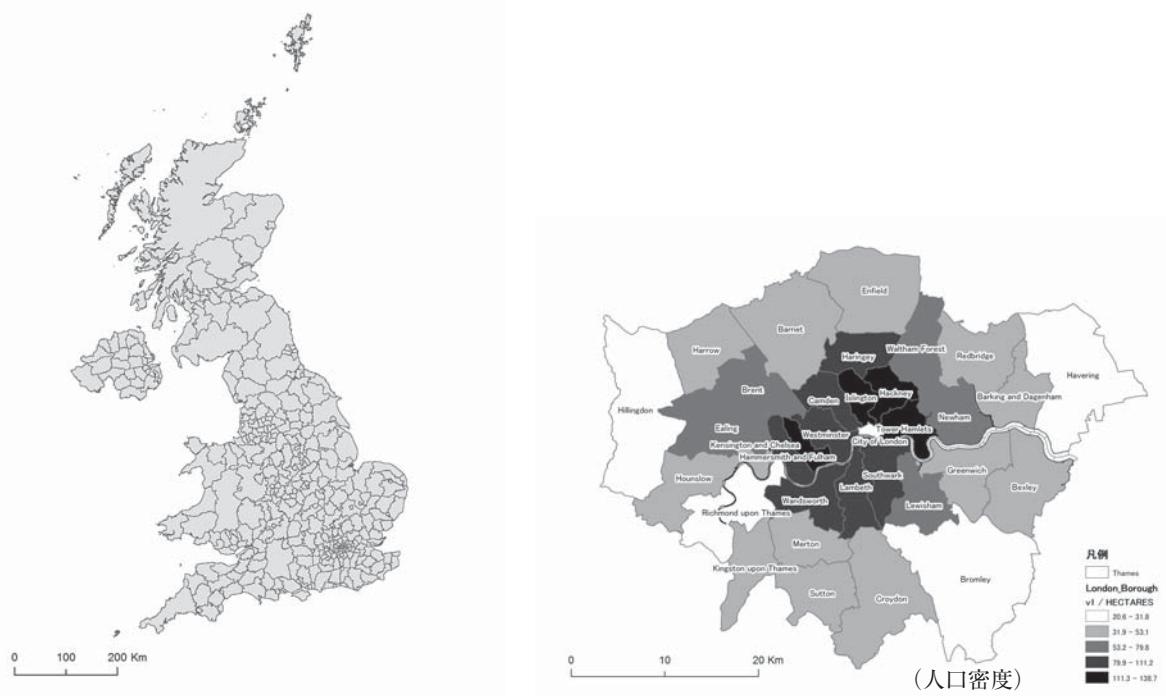


b) Regions (イギリスで9)

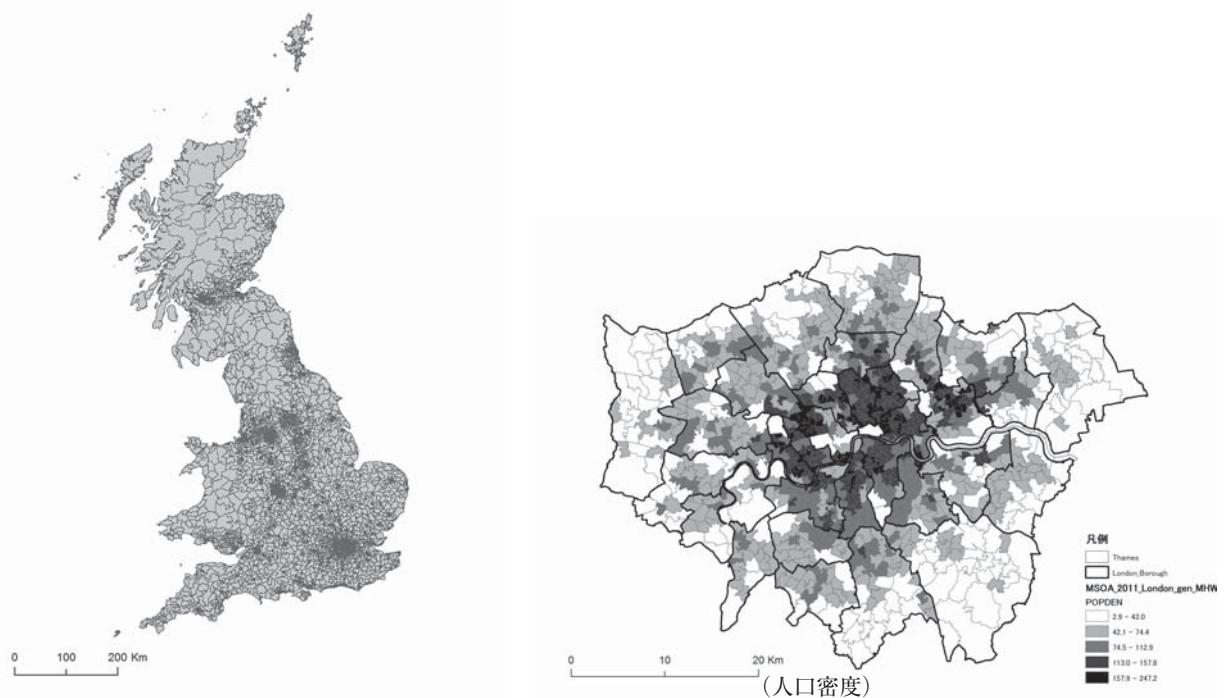


c) Counties (イギリスで55)

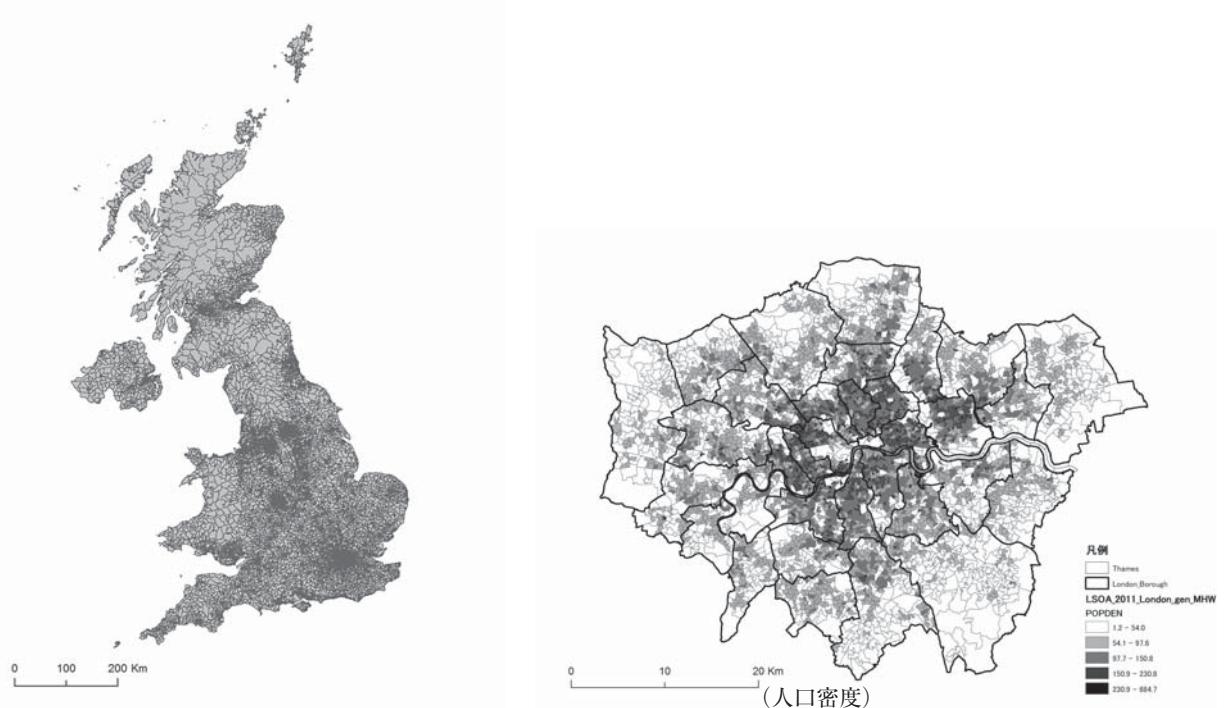
第7図 英国の2011年国勢調査の空間単位



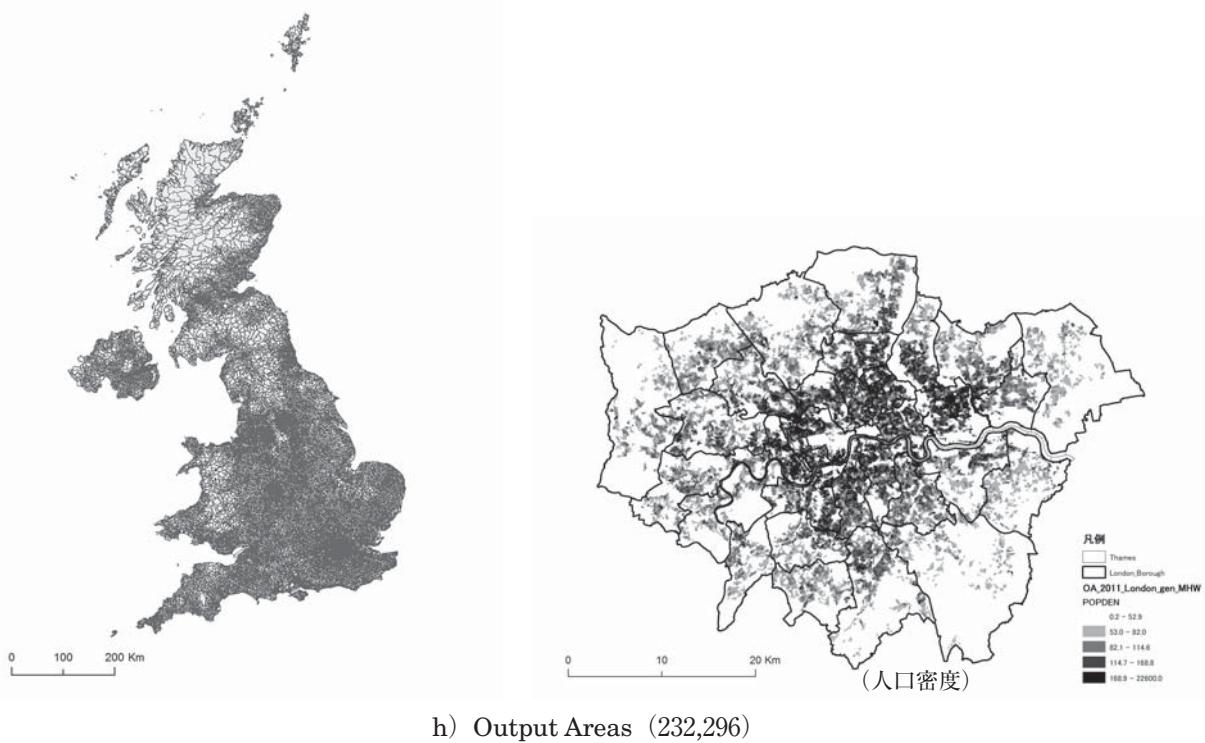
第7図 (続き)



f) Middle Layer Super Output Areas and Intermediate Zones (8,436)



g) Lower Layer Super Output Areas and Data Zones (42,143)



第7図 (続き)

公開を行っているが、英国の国勢調査の公開は、前述のように ESRC の基金による UK Data Service によって行われている。UK Data Service は、研究・教育目的のためにあらゆるデータの利用ニーズを満たすために設立されたもので、そのデータは、国勢調査以外にも、英国政府による様々な統計データを含んでいる。UK Data Service のセンサス・サポートは、イングランドとウェールズ (ONS)、スコットランド (NRS)、北アイルランド (NISRA) の 3 つのセンサス局から国勢調査の元データの提供を受け、英国情報システム合同委員会 (JISC)、UCL (University College London)、Edinburgh 大学、Leeds 大学、Manchester 大学の地理学を中心とするの研究者集団が、統計表や GIS データのオープンデータを作成し、公開している。英国では、国勢調査を研究する研究者によって、公開システムが構築されているために、研究者が教育・研究に利用しやすい構造になっている。

一方、日本の場合は、国が主導となって、急速にオープンデータ化を進めたために、必ずしも、研究者の利用しやすい形にデータが公開されていない。その背景には、これまで日本の国勢調査の属性データと境域データが、Sinfonica から有償で頒布されてきたことにも一因があると思われる。例えば、民間企業は、Sinfonica からそれら国勢調査データを加工して販売する権利を含めた第三者提供価格でデータを購入し、GIS で利用しやすい形で販売してきた。そして、研究者は、教育・研究用に、それらのデータを Sinfonica から民間企業と同様に購入するか、第三者提供の権利をもった民間企業が作成したデータを購入してきた。いずれにせよ、国勢調査データは高額で、研究者は研究のたびにこれら統計データを購入する必要があった。このような状況は、基本的に英国でも同様であるが、英国では、研究費を配分する ESRC が、研究者の国勢調査利用を一元化して、基本的に、登録した研究者のみがアクセス可能な Casweb のような Web サイトを構築してきたのである。そして、そのような登録された研究者のみが利用可能であった国勢調査の公開システムが、オープンデータ

の展開の中で、UK Data Serviceにおいて、研究者のみならず民間も含めた社会全体へ公開するシステムへと移行したといえる。もちろん、こうした状況の変化は、国勢調査の小地域統計などでは、データ量が膨大となることから、インターネット環境やデータ貯蔵技術の発展とも大きく関連している。

第2章で述べたように、e-Stat から公開されている 2010（平成 22）年国勢調査以前の統計表は、Sinfonica が頒布している全ての統計表に対応していない。それゆえ、総務省統計局が提供する統計表の完全な属性データは、Sinfonica が頒布しているものを利用せざるを得ない状況にある。これに対して、日本の研究者への国勢調査などの官庁統計の提供は、全国共同利用・研究拠点である東京大学空間情報科学研究センター（CSIS）がその役割を果たしている。CSIS では、「CSIS 空間データの利用を伴う共同研究（JoRAS）」¹⁸⁾ の中で、Sinfonica が頒布する国勢調査を含めた多くの官庁統計の研究利用を推進している。

また、民間では、前述のように、これまで sinfonica から第三者提供価格で国勢調査の統計表と境域データを購入し、GIS ソフトで即座に利用可能な形式に加工するなどしてそれらを販売してきたが、オープンデータの展開の中で e-Stat に公開された国勢調査の属性データと境域データを活用して、無償でデータ加工を行いそれを販売している。例えば、ESRI ジャパン（株）の「スタートパック（基本統計）」¹⁹⁾ のように、世界標準の GIS ソフトである ArcGIS で即座に利用可能なデータセットを開発・販売することで、新たなビジネス展開を図っている。これらの製品は、「政府統計の総合窓口（e-Stat）利用規約」に準じて、e-Stat からダウンロードしたデータは、出典を記載すれば、第三者に提供することを目的としても自由に利用することができるという規定の下で、開発・販売されている²⁰⁾。

また、日本では、新統計法においては、個人情報保護の下でのオーダーメード集計に加え、インターネットを介したオンデマンド集計の利用条件の緩和が進められており、これまで総務省統計局が業務的に作成してきた統計表以外の新たな統計表の作成が期待されている²¹⁾。これに対して、英国では、日本のオーダーメード集計に対応するものとして、Commissioned tables がある。これは、通常の統計表以外に、自治体などの依頼によって作成された特別な統計表で、個人情報の保護を厳守したうえで、作成後は Web に公開され誰もが利用できる形で公開されている²²⁾。

UK Data Service のセンサス・サポートでは、本稿で紹介した集計データ以外に、人口流動データやマイクロデータなども提供されている。さらに、研究者や民間によって国勢調査を用いて作成された、Population data、Deprivation data、Census area classifications なども Census related として公開されている²³⁾。これらは、膨大な国勢調査のデータを用いて研究あるいは商業ベースで作成されたもので、剥奪指標や、小地域に基づいた地区類型であるジオデモグラフィックスが、産官学連携の下で UK Data Service のセンサス・サポートから公開されている。

このように英国では、国勢調査を用いた研究・教育を推進してきた研究者が、国勢調査の統計表や境域データの作成や公開に関わることによって、研究者目線での利用しやすいオープンデータ化が実現されていることは注目に値する。こうした先進的な点は、日本の国勢調査のオープンデータ化にも積極的に取り入れていくべきであろう。このような地理空間情報のオープンデータ化が、地理情報科学、ひいては地理学のさらなる発展には不可欠である。

5. おわりに

ここ数年のオープンデータの推進の潮流の中で、国の最も基本的な統計である国勢調査が Web を介して大量に公開され始めた。こうした動向は、日本だけでなく国勢調査を用いた地理学的研究が盛んな英国においても同様で、これまで研究者への研究・教育目的に限定されてきた Web システムが、一般の利用も含めて広く公開されるようになった。日本でも政府統計の総合窓口である e-Stat を介して、多くの統計表が公開され、それに対応する境域データも提供されるようになった。本研究では、日本と英国の国勢調査のデータ公開方法を紹介し、その違いに着目した。そして、今後の日本の国勢調査のオープンデータ化に際して、英国で展開するような、産官学連携での国勢調査のデータ公開や活用が参考になることを示唆した。

英国では、オープンデータ化とともに、現在、ビッグデータを用いた研究が注目を集めている。例えば、UCL、Leeds 大学、Liverpool 大学、Oxford 大学などが中心となって展開する CDRC (Consumer Data Research Centre)²⁴⁾ は、ESRC の多額な研究費によって、民間の消費者行動に関する膨大なジオ・ビッグデータを集積させて、産官学での研究を展開している。こうした秘匿性の高い消費者データと、オープンデータ化された国のデータを融合させた新たな地理情報科学の研究が展開されている。日本においても、2016 年度から一般社団法人社会基盤情報流通推進協議会のもとに、G 空間情報センター²⁵⁾ が設立された。そこでは、国・自治体そして民間で作成された膨大な地理空間情報のポータルサイトを構築するものである。この設立は、2012 年 3 月に政府で閣議決定された地理空間情報活用推進基本計画に基づいて、G 空間社会の実現に向けた展開の 1 つである。このような産官学連携によるオープンデータ化の推進を今後さらに高度化させていく必要がある。

付記：本研究の一部は、科学研究費補助金（24320171 および 15H02964）の助成を受けた。また、立命館大学の 2016 年度前期学外研究において、4 月から 7 月に滞在した UCL の CASA (Centre for Advanced Spatial Analysis) での研究成果の一部でもある。滞在期間中お世話になった、CASA の Andy Hudson-Smith 教授・所長と、地理学部の Paul Longley 教授に感謝する。

注

- 1) Goodchild, M. F.: Geographical information science. *International Journal of Geographical Information Systems* 6-1, 1992, pp.31-45.
- 2) Paul A. Longley, P. A., Goodchild, M. F., Maguire, D. J. and Rhind, D. W.: *Geographic Information Science and Systems*, 4th Edition. Wiley, 2015.
- 3) 日本学術会議地域研究委員会・地球惑星科学委員会合同地理教育分科会『提言：地理教育におけるオープンデータの利活用と地図力/GIS 技能の育成—地域の課題を分析し地域づくりに参画する人材育成—』、日本学術会議、2014。
- 4) 昭和の大合併により市町村数は、1950（昭和 25）年時点の 10,500 から、1960（昭和 35）年時点の 3,574 へと大幅に減少した。
- 5) 羽渕達志（2012）：「国勢調査の地域区分と地域データについて」、2012、
https://www.nstac.go.jp/services/society_paper/24_06_01_2Paper.pdf (2016 年 11 月 30 日閲覧)。
- 6) 浅見泰司・矢野桂司・貞広幸雄・湯田ミノリ編『地理情報科学—GIS スタンダード』、古今書院、2015。
- 7) 「政府統計の総合窓口（e-Stat）」 <https://www.e-stat.go.jp/estat/guide/index.html> (2016 年 11 月 30 日閲覧)。
- 8) 国土数値情報ダウンロードサービス <http://nlftp.mlit.go.jp/ksj/> (2016 年 11 月 30 日閲覧)。

- 9) Municipality Map Maker for Web 市区町村区域の GIS データ生成ツール <http://www.tkirimura.com/mmm/> では、1970 年 1 月 1 日から 2016 年 10 月 9 日まで対応している（2016 年 11 月 30 日閲覧）。桐村喬・中谷友樹・矢野桂司「市区町村の区域に関する時空間的な地理情報データベースの開発－Municipality Map Maker for Web－」、GIS－理論と応用 19-2、2011、83-92 頁。
- 10) 明治期に天文観測で決定した、東京天文台の位置を日本経緯度原点とした「日本測地系」が、超長基線電波干渉法（VLBI）や人工衛星を用いた観測に基づく地球全体によく適合した測地基準系である世界測地系に移行した。
- 11) ESRI ジャパン「標準地域メッシュが簡単に作成できる便利なツール」
<https://blog.esrij.com/2016/03/15/post-e331/>（2016 年 11 月 30 日閲覧）。
- 12) 梶田真：「国勢調査における小地域統計の整備過程とその利用可能性」、東京大学人文地理学研究 13、2008、31-43 頁。
- 13) これらの各センサス局からも、2011 年、2001 年センサスの全ての統計表がホームページから公開されている。例えば、イングランドとウェールズの ONS からは Nomis のサイト (<https://www.nomisweb.co.uk/>) から様々な統計表をダウンロードできる。
- 14) Rhind, D.: *A census user's handbook*. Methuen, 1983. Dale A. and Marsh, C. eds.: *The 1991 Census User's Guide*. HMSO, 1993. Openshaw, S. ed.: *Census Users' Handbook*. Wiley, 1995.
- 15) 磯田弦：「デジタルデータ供給機関の展開－イギリスからの二つの寓話」、じんもんこん 2005 論文集、2005、225-232 頁。
- 16) UK Data Service <https://census.ukdataservice.ac.uk/>（2016 年 11 月 30 日閲覧）。
- 17) David, M.: *The UK Census of Population 1991*. Sage, 1993.
- 18) 「CSIS 空間データの利用を伴う共同研究（JoRAS）」http://www.csis.u-tokyo.ac.jp/japanese/research_activities/joint-research.html（2016 年 11 月 30 日閲覧）。
- 19) ESRI ジャパン（株）の「スタートパック（基本統計）」<http://www.esrij.com/products/data-content-starter>（2016 年 11 月 30 日閲覧）。
- 20) 政府統計の総合窓口（e-Stat）利用規約 <http://www.e-stat.go.jp/estat/html/spec.html>（2016 年 11 月 30 日閲覧）。
- 21) 総務省「公的統計の整備に関する基本的な計画」の実施状況とデータ提供の新たな取組について
<http://www.stat.go.jp/training/2kenkyu/pdf/gakkai/toukei/2015/uematsu.pdf>（2016 年 11 月 30 日閲覧）。
- 22) 英国 2011 年国勢調査の Commissioned-tables は以下から、
<http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/census/2011/census-data/2011-census-data-catalogue/commissioned-tables/index.html>（2016 年 11 月 30 日閲覧）。
 そして、2001 年国勢調査のものは以下から閲覧できる。
<http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/census/census-2001/data-and-products/data-and-product-catalogue/commissioned-output/index.html>（2016 年 11 月 30 日閲覧）。
- 23) Census Related Data <https://census.ukdataservice.ac.uk/get-data/related>（2016 年 11 月 30 日閲覧）。
- 24) Consumer Data Research Centre <https://cdrc.ac.uk>（2016 年 11 月 30 日閲覧）。
- 25) G 空間情報センター https://www.geospatial.jp/gp_front/（2016 年 11 月 30 日閲覧）。

(本学文学部教授)

Current Status and Issues of Data Disclosure Methods of GIS-based Population Censuses:
Through the Comparison between Japan and the United Kingdom

by
Keiji Yano

This paper aims to consider how, in the midst of the recent trend towards disclosing open data of the state or local governments, data of population censuses based on the geographical information system (GIS), which are a fundamental source of national geographical space information, are being disclosed in Japan and the United Kingdom. After comparing the two countries, this paper discusses some future perspectives on the disclosure method of population censuses in Japan and the direction of census geography.

The Japanese population census started in 1920. Since the period of high economic growth, information on spatial units smaller than municipalities have been collected using regional mesh statistics and the like, and, since 1990, data improvement has been carried out using the GIS, which follows basic unit areas. Meanwhile, in England, starting with the 1961 population census, spatial units smaller than Wards called EDs (Enumeration Districts) were introduced, and, since the 1971 population census, statistical tables for each ED etc. have been disclosed along with GIS data.

In the United Kingdom and the United States, which are developed countries in terms of GIS utilization, numerous census geography handbooks for utilizing population census data for geographical research have been published. These handbooks not only contain data provided by the government, but also point to tasks and issues concerning the development of analyzing methods using the GIS. Furthermore, by processing official statistical data, these handbooks add new value to geo-spatial information. The present paper compares the way that population censuses are disclosed in Japan and the England, and makes some suggestions as to the possibility of a unique development of census geography in Japan.