

オケーショナル・ペーパー No.26

ジオコード情報の活用による統計の 把握精度改善の試み

アメリカセンサス局 2010 年センサス改善計画を事例として

2011年9月

法政大学

日本統計研究所

ジオコード情報の活用による統計の把握精度改善の試み

アメリカセンサス局 2010 年センサス改善計画を事例として

森 博美

はじめに

近代統計が時点と並んで統計単位の存在そのものを規定する二要因の一つである場所をどう捉えてきたかについては、すでに別稿〔森(3)〕で論じた。個票調査の導入以来、これまでの統計学では統計単位が統計調査による直接的把握の対象とされてきた。しかしそこでは、調査があくまでも安定的な集計量を得るための統計単位情報を提供するものという観念、言い換えれば「統計＝集計量」という統計観がこれまでは支配的であった。さらに、センサスのように母集団を構成する全ての要素を悉皆的に調べる調査によって正確な結果数字を獲得するためには、統計単位は、漏れることなくしかも重複なくまさに 1 回限り調査されなければならない。このために、統計単位が存在する空間的な範囲を相互に排反的な区域(調査区)群によって網羅的にカバーすることが調査実施上の要件として求められてきた。

これまでの統計調査は、個々の統計単位が本来的に持つ場所情報を、個々の調査区の場所情報、すなわち調査区コードとして捉えてきた。これには二つの事情、すなわち、調査実施上の要件としての調査区の設定という事情、それに、統計を第一義的に集計量とする統計観が場所情報を一定の面的広がりにおいて捉えてきたという事情がある。

調査区コードは、当該区域内に存在する個々の統計単位が持つ場所情報を、調査区という一定の面的広がりにおいて集約的に表現したものである。一般に同じ調査区に属する複数の単位に対して同一の調査区コードが付与される。このため、単一の調査区内では、場所に関して個々の単位に帰属する位置情報と調査区コードの間には、いわばn対1の関係が成立している。すなわち、把握漏れと重複調査の排除という調査技術上の必要から設定された調査区が、調査個票に記載された調査区コードを介して調査結果が持ちうる情報量に影響を及ぼしている。すなわち、このような位置情報に関する取り扱いが統計単位が保有する統計単位情報と調査票に転写されるそれとの間の情報のギャップ、少なくとも場所に関する情報喪失を発生させているのである。

筆者は、統計調査によって把握される統計単位情報(変数群)を、data carrier と data body とに範疇的に区別することで、統計個票情報の情報特性の解明を行った〔森(1)〕。調査個票情報のうち個体の属性情報、調査項目、さらには調査員記入事項として記載された諸事項に関しては、量的変数だけでなく質的変数情報についても、多様な集計結果表が作成され、統計として提供されてきた。また、調査区コード、都道府県コードや市区町村コードといったコード情報は、地域別集計結果表の作成の際に集計キー変数として使用されてきた。また、国勢調査などでは調査区あるいは基本単位区別の調査結果は、必要に応じて同定作業を行うことで、地域メッシュ統計や統計GISといった統計の新たな表章形態あるいは統計の利用面で活用されてきた^(注1)。

一方、同じく位置情報である世帯や事業所等の統計単位の存在の場に関する住所(あるいは所在地)のこれまでの統計上の取り扱いは、上記の場所コードのそれとは異なる。なぜなら、それらは機械的あるいは目視による照合情報の一つとして用いられることはあっても、それを数値情報化し

て統計作成業務に活用することは行われてこなかったからである。

全地球測位システム(Global Positioning System: GPS)は、もともとは軍事情報技術として開発されたものであった。近年のその著しい精度の向上、さらには、対応ソフトを装備した情報端末の急速な低廉化により、民生分野においてこの技術を用いた様々なサービスが提供されるようになってきている。

GPS コードを調査単位の住所(所在地)と関連づけることで、調査区という面情報とではなく世帯や事業所といった各個体の所在の場を示すポイント^(注2)情報として表示することができるようになる。調査単位から得られる統計単位情報を記録した個体レコードに新たに GPS コード情報を変数として付加し GPS 付レコード(GPSed records)とすることで、位置情報に関して調査区コードという形での位置情報の付与が不可避免的に伴っていた本来の調査個票情報からの情報喪失を回避することができるように思われる。

本稿では、GPS コードの統計の作成面への活用の一つの具体的な先進事例として、合衆国センサス局が 2000 年センサス以降「2010 年センサス準備計画」(2010 Census Re-Engineering Plan)の一環としておよそ 10 年の歳月をかけて進めてきた「MAF/TIGER 改善計画」(MAF/TIGER Enhancement Program :MTEP)を取り上げ、GPS 情報技術の活用によるセンサスの把握度改善について紹介する。

1. MAF/TIGER Enhancement Program(MTEP)について

2002 年 11 月 22 日に東京で開催された「人口センサスとセンサスマイクロデータに関するワークショップ」において、当時の米国センサス局人口センサス部副部長 Waite, P.J.氏から「米国の 2010 年センサスの方法と展望」という報告^(注3)が行われた。Waite 報告によれば、「2010 年センサス準備計画」は、センサスの品質改善に向けた 3 つの相互に連携した内容からなる。アメリカ地域社会調査(American Community Survey: ACS)の導入、それに伴う詳細調査票(long form)の人口センサスからの分離による基礎調査票(short form)に特化した調査実施計画の実施^(注4)、それに MAF/TIGER 改善計画(MTEP)がそれである。

Waite 報告によれば、MTEP は、2000 年センサスで使用したマスター住所ファイル(Master Address File: MAF)と電子地図(Topologically Integrated Geographic Encoding and Referencing: TIGER)を地球測位系座標に調整し、TIGER データベースを新たに市販のデータベース環境にも対応できるようにすることを主たる内容とするものであった[(7) p.25]。報告はまた、更新された住所情報を持つ MAF と GPS による位置確認を行った TIGER データベースが、住所や構造物のリストとその位置をより包括的にしかも高い精度で把握することをセンサス局に可能にするものであり、そのためには GPS 機能付のハンドヘルド PC の全面活用が不可欠であるとしている[(7) pp.25-26]。このような最新技術を用いたより精度の高いセンサス結果を提供することは、センサスに対する合衆国憲法による要請でもある。

以下に、2000 年人口センサスの精度に関する結果評価とそこでの課題、GPS 情報技術の利用によるセンサスの制度改善策、さらには具体的な MTEP の取組み内容について検討する。

2. 2000 年センサスの結果精度評価

センサス局では、センサスの精度について、これまで郵送回答率 (mail response rate) によってその評価を行ってきた。ここでの郵送回答率とは、米国郵政公社 (U.S. Postal Service) による郵送で「あて先不明 (undeliverable)」により返送されてきたその存在が確認できない住戸や非居住の住戸も含む調査票の送付総数に対する郵送による回答調査票の割合によって示されるところの郵送による調査実施の結果精度の評価指標である。センサス局では2010年センサスから新たな評価尺度として、郵送報告率 (mail return rate) あるいは郵送参加率 (mail participation rate) を採用した。これは、発送総数から「あて先不明」として郵政公社に返送されてきた調査票を除いたものを分母として、それに対する郵送あるいは調査員による手渡しにより居住者のいる住戸に届けられた調査票数の割合を示すものである〔(5) p.3 注 1、(45) p.2〕。その結果、郵送報告率 (郵送参加率) は、当然ながら「あて先不明」として郵政公社に返送の分だけ郵送回答率よりも高めに出ることになる。

1970年センサスで郵送調査 (郵送による配布・回収 (mail out/mail back)) が全面的に採用されて以来、87% (1970年) → 83% (1980年) → 74% (1990年) と次第に低下してきた郵送報告率は、2000年調査でも 72% と、その低落傾向に歯止めをかけることはできなかった〔(5) p.12〕。また、回答率 (response rates) も 78% (1970年) → 75% (1980年) → 66% (1990年) → 66% (2000年)、と近年やや下げ止まりを示しつつも、漸減傾向にある。また州別に見ても、2000年センサスでは郵送報告率が1990年のそれを上回ったのは、マサチューセッツ州などわずか3州だけで、残りの47州ではいずれも低下している〔(5) p.13〕。なお、郵送回答率の面では、1990年調査時の (65%) とほぼ肩を並べる64%が2000年調査では得られたが、調査票の回収には前回の調査に比べて3.5倍以上の経費〔注5〕を要したと言われている〔(5) p.3〕。

調査員が戸別訪問により未回答者に対する記入依頼を行う督促調査は、例えば2000年センサスの中で最も労働集約的作業とされており、それに対して支出された経費約14億ドルは、1991会計年度から2003年度までの2000年センサス関係予算総額65億ドルの実に22%を占めている〔(5) p.1〕。なお、2000年調査では、未回答者の督促のために調査員が訪問した世帯のうち773,000世帯では、すでに調査票は提出済みであり、それに要した費用は2200万ドルにのぼる〔(5) p.3〕。

2000年センサスは、その実施担当者自身による評価では、「実施状況並びにカバレッジの両面でこれまでのセンサスの中で最も成功裡に行われた」〔(7) p.30〕とされている。しかしながら、会計検査院 (GAO) による同調査に対する評価は、これとはやや異なった内容のものである。ちなみに、『報告書』(2000 Census - Best Practices and Lessons Learned for More Cost-Effective Nonresponse Follow-up) (2002年2月刊行) は、郵送法がセンサスに本格的に導入されて以降のセンサス実施に対する協力度の動向も含め、2000年センサスの結果精度等の改善に向けて、次のような分析と勧告を行っている。

このような2000年センサスの実施状況評価に基づき、会計検査院では、商務省に対してセンサス局が未提出者の督促に可能な限り費用対効果の面で有効な政策を導入するよう要請している〔(5) p.33〕。

3. MAF/TIGER とそれらが抱える問題

(1) MAF

MAFは、1990年センサスの後に米国郵便公社(U.S.Postal Service: USPS)が保有する配達住所リスト(Delivery Sequence Files)を主要な情報源とし、それに州、郡、市といった地方政府が持つ集団施設^{〔注6〕}等に関する所在地情報、さらには調査員による現地踏査から得られた様々な副次的情報を統合することによってセンサス局が独自に構築した住所データベースである。それには、アメリカ50州並びにプエルトリコ、米統治下の島嶼地域に所在する1億1500万の住戸および6000万のビジネス組織等の所在地情報が格納されている〔(6) p.(i)、(37)〕。

MAFは二種類の住所情報、すなわち調査票郵送用の宛名情報と未回答者に対して調査員が訪問回収を行うための現住所情報を持っている〔(36) p.2〕。なお、MAFがこのような個人情報からなるデータベースであることから、合衆国法典第13巻の規定により、その使用はセンサスの実施に限定されている〔(6) p.1〕。

(2) TIGER

センサス局は、1960年代末に主要都市地域を対象に、センサスブロックの端点を直線で結んだ多角形(ポリゴン)からなる図形データベースとしてDual Independent Map Encoding (DIME)を構築した。しかしながら、このデータベースは、国土の2%弱しかカバーしておらず、また地域別にそれぞれ独立したファイルとなっており、全地域を1枚の地図としてシームレス的に表示できるものでもなかった。

1983年にセンサス局にデジタル調査地図の開発予算が認められた〔(7) p.1〕。それを受けて同局は、米国地質調査局(U.S. Geological Survey: USGS)の10万分の1地形図を用いて水路、鉄道その他の地勢情報をラスター型データとしてスキャンし、道路のセンターラインデジタル情報、航空測量図、それに自治体が保有する地理情報システム等を用いた地図データベースとして独自に開発したのがTIGERである。それは、都市型居住区域に関して342区分からなる地図を持ち、全米50州の他、プエルトリコ、さらに米国統治下にある島嶼部をカバーするものであった〔(37) p.5〕。なお、システムの構築に当たっては、当時のTIGERには、調査員の手書きによる地図をデジタルスキャンした情報も一部含まれていた〔(7) p.1〕。また地図情報の更新は、職員が米国地質調査局の2万5000分の1の地図を手作業で数値化することで行われた〔(47) p.2〕。

データベースとしてのTIGERは、道路、建物、河川、湖沼の情報を持っている。また、TIGERによって表示される地域の中には、州、部族地域、都市、郡、国会議員選挙区、学校区といった政治・行政的境界を表現するものや、都市統計区(Metropolitan Statistical Areas :MSA)、センサス調査区、センサスブロックグループ、センサスブロックといった統計調査区分を示すものもある〔(47) p.2〕。この他にもZIPコードによる集計地域(ZIP Code Tabulation Areas: ZCTA)^{〔注7〕}がある。TIGERの保有情報は都市部とそれ以外の地域とで異なり、都市部については、街路の名称と位置、河川、鉄道、境界、その他の地勢情報、それにMAF住所情報を持っている。他方、非都市部については、呼称を持たない道路も存在する。さらに非都市部についてTIGERは、住戸の位置情報は持っているが住所情報は保有していなかった〔(36) p.2〕。

TIGERは、実査過程での調査員の調査区における活動支援や調査結果の表章において使用されるが、このデータベースを全面的に使用できるのはセンサス局に限定される。なお、TIGERが持つ情報の一部から編成されたTIGER/Lineファイルは、これまでも州、地方、部族政府、民間企業、教育機関等に提供され広く使用されてきた〔(6) p.2〕。

TIGERは1990年センサスで初めて使用された。しかしその後、センサス局には境界座標の更新や構造物の改廃情報を恒常的に更新するだけの予算措置がなかった。また、センサス局が内生的

に開発した TIGER は、いくつかの問題点をかかえていた。すなわち、①マルチユーザーに対応できないこと、②TIGER の使用者の養成・訓練に多大な時間が必要であったこと、③ウェブ経由でのデータベースへのアクセスに支障をきたしていたこと、④データベースの拡張、修正の自由度に欠けていたこと、⑤COTS(商用品 Commercial off the shelf)や GIS を使用した方がデータの共有が容易であること、⑥郡(county)別のシステムから全国ベースのシステムへの移行がシームレス的でなかったこと、などがそれである〔(8) p.1〕。また、その後 TIGER には度重なる手直しや補正が加えられ、最終的には、TIGER は 3300 を超えるステップ数を持つ巨大なプログラムへと肥大していた〔(6) p.2〕。なお、商務省査察局は、センサス局によるこういったソフト開発に対して、「アドホック的で適正な管理、試験が欠落する傾向がある」〔(6) p.(ii)〕といった苦言を呈している。

この他にも TIGER は、保有する地図情報の精度の点でも大きな問題点を持っていた。道路や水路はしばしば実際とは異なる場所に描かれ^{〔図1〕}ており、同じ地図上でもその乖離度が 10m のところもあれば 100m 以上ずれているところがある^{〔図2〕}など、地図の精度は著しく不均質なものであった。

4. MTEP の目的

米国政府会計検査院(GAO)の指摘を受けてセンサス局が掲げた 2010 年センサスに向けての改善計画は、作業リスクの削減やカバレッジの向上を主たる内容としたものであった。しかもそれらをより少ない経費によって達成することが同局には求められていた〔(6) p.1〕。また、2002 年に議会は、承認した延べ 2.1 億ドルの予算によって、1980 年代に内生的に開発した TIGER の商用データベースへの対応可能なものへの再編を軸とする MAF/TIGER の改善をセンサス局に対して命じた。これを受けてセンサス局では、直ちに再設計のための委員会として MAF/TIGER Redesign Steering Committee を組織し、以下のような内容を持つ MAF/TIGER 改善計画(MTEP)の策定に着手した。

MTEP では、センサスにおける調査員による実査業務の GPS 機能付のハンドヘルド PC による支援、地方政府による TIGER データベースへのウェブ経由でのアクセス可能性の保証、操作上の諸要請に対するより素早い対応、データベースの更新作業の容易化、MAF/TIGER 作業員の訓練期間の短縮を可能にする措置、などが求められた。より具体的内容としてそれは、実査において調査員が使用する PC への住所情報、地図情報の供給が可能であること、また約 39,000 の地方政府や郵政公社等からのデータベースのバッチ方式による同時更新が可能であることなどを含んでいた〔(6) p.3〕。

センサス局では、具体的なデータベースシステムの再設計に当たって、①広範に使用されている専用のデータベースソフト ORACLE の導入、②より多様なアプリケーションの活用が可能な商用ソフト(COTS)と GIS が使用できること、③ソフト工学研究所の Capability Maturity Model(CMM)のレベル2に対応すること、④ニーズの変化への対応を考慮してオープンデザイン型のソフトとすること、⑤MAF と TIGER 一体型のデータベース仕様とすること、⑥データベースの修正や更新が容易にできるように、その機能面での改善をはかること、などに留意することが確認された〔(8) p.1〕。

5. MTAIP について

(1) TIGER の再設計と MTAIP

MTEP は、MAF と TIGER とのデータベースとしての実質的統合を目標とするものであり、地図情報の精度の向上がその前提となる。MTEP のまさに情報面での中核部分を形成しているのが、MAF/TIGER Accuracy Improvement Project (MTAIP)である。

MTAIP においては、部族政府、州政府、さらには郡あるいは市といった様々なレベルの自治体政府その他の小地域に関する結果表章の精度を向上させることで、選挙区の区割りや地域区分の変更の際には連邦、州、地方政府に貢献することが謳われていた〔(14) p.11〕。そのためにも、TIGER の精度向上と商用ソフト(COTS)や GIS への対応が可能でしかも膨大な情報の日常的な更新機能を持つデータベースへとシステムの抜本的に再構築することが求められた。なお、MTAIP に関係した業務そのものは、2006 年までには終了した。

このような、データベースそのものに求められる新機能に加え、新たに構築するシステムには、充足すべき重要な制度的要件として、それが調査実施フレームとしての機能を果たせるように、何よりも使用料あるいは著作権による使用制限が発生しないことが求められた〔(35) p.1〕。MAF と統合された TIGER が変数として保有する個々の GPS 情報は、合衆国法典第 13 巻第 9 編^{〔注8〕}によってその秘密が保護される〔(41) p.2〕。MTAIP によって構築される新たな TIGER は同法による保護の対象とされている情報を含むものではないとの認識にセンサス局は立った。

センサス局では、TIGER を社会的に汎用性のあるソフトとして、政府だけでなく教育機関や民間を含めて広範な利用者に提供するために、道路のセンターライン、通りの名称、水域の位置、境界といった全国土空間情報基盤(National Spatial Data Infrastructure: NSDI)が有する精度の高い空間データの使用制限を容認することはできないという立場からシステムの再設計にあたった。このような目的に照らして、政府による NSDI と民間部門による種々の付加価値情報の一体利用という状況を一種の理想型として想定し、それに利用制限を設けることなく、民間による情報の付加を許容するシステムとするという設計方針を持っていた。そこでは、センサスブロックのような小地域情報の利用についても特別な課金を発生させることなく、多様な利活用が可能な情報インフラとして整備することを目的としていた。

(2) 地図の精度改善

センサス局では 2002 年から ArcPad ソフトを試験的に使用〔(26) p.2〕してきており、MTAIP でも空間情報の取得に GPS が積極的に用いられた。センサス局にとっての GPS 使用のメリットとしては、紙ベースでの調査区地図が不要となること^{〔注9〕}、デジタル情報として通りや住戸の所在地の名称が事前入力されているため更新業務が大幅に削減可能であること、住所情報がなくとも調査員は調査客体の住戸を特定できることなどがあげられる〔(37) p.34〕。さらに、GPS 情報を用いて住戸の空間的位置を確定することで、MAF に時として含まれる住所情報の重複登録を検出することができる^{〔図 3〕}。この他にも、空間情報の取得に GPS を使用することで、地方自治体等の公的部門や民間部門が保有するデータとの互換性が担保されるというメリットもある。

このうち地図ならびに空間情報の精度向上については、従来の TIGER が座標精度 50m であったものを CE95 7.6m^{〔注10〕}へとその精度の大幅な向上が図られることになった。TIGER の位置情報の改善のために、MTAIP の作業過程でセンサス局の地方事務所は、部族政府、州、郡、地方といった各レベルの地方自治体に協力を求め、それらが保有する GIS ファイルの入手を行った。精度面で上記の基準に満たないファイルについては、センサス局から業務を受託した Harris 社が商用 GIS 等から情報の入手を行なった。また精度要件に適合した地図情報が存在しない郡については、同社が独自にセンターライン情報の取得を行った〔(14) p.14〕。

センサス局地理部 (Geography Division) からその業務を受注^[注 11]した Michael Baker Jr. 社 (Michael Baker 社の子会社) では、TIGER の地図情報の精度向上のために GPS 情報を提供するとともに、2007 年までに全米 3,232 の郡のうち 2,865 の郡で、また 2008 年中に残りの郡についてもセンターライン情報の再取得を完了した [(16) p.2]。道路のセンターライン情報の整備によるより高精度の地図情報の獲得は、その後予定されている address canvassing による住戸等の位置情報との整合性を持たせる上で不可欠のものである。なお、[図 4] は、センターライン情報の精度の向上による地図の改善を対比的に示したものである。

(3) センサス住所の地域更新作業

2010 年センサス実施年の前年に行われた住所の最終確認作業 (address canvassing) に先行してセンサス住所の地域更新作業 (Local Update of Census Addresses: LUCA) が行われた。

2006 年度の MTAIP の完了を受けて、2007 年秋から約 120 日間かけてセンサス局が保有するセンサスの住所情報の地域更新作業が各自治体の協力を得て行われた [(16) p.3]。11,500 以上の各地方自治体その他の組織からの協力、支援の下に、連邦政府だけでなく部族政府、州、郡、地方といった地方自治体、さらには民間部門に対しても開かれた共通の情報プラットフォームの構築を目指すこのプロジェクトは遂行された [(46) p.1]。特に地方自治体から提供された情報は、新設の道路や建物といった新たな構造物あるいはその所在地情報の有力な源泉となった [(37) p.3]。この他にも、高価な有料ソフトに代わり無料で政府から提供されることになるデジタル地図の整備のために、民間企業^[注 12]もその構築にあたって地理部への積極的支援を行った [(37) p.6]。

センサス局は質の高い GIS 情報の提供をこういった政府等から著作権や使用料の制限なく受け取ることで、更新ファイルの TIGER Enhancement Database (TED) と呼ばれる中央リポジトリへの格納が行われた [(35) p.1、(37) p.4]。このようにして TED に格納された住戸等の所在情報リストはセンサス局から各自治体に提供され、その位置情報等の確認作業が行われた [(32) p.6]。このような一連の作業の結果、より精度の高い地図情報を持つ TIGER をセンサス局は構築することができた。この他にも地理部では、全米地質調査局 (National Geological Service: NGS) の協力を得て、address canvassing によって取得した GPS 座標情報の補正システムを構築した [(50) p.4]。この補正プログラムは、数度にわたる検証テストを経て、最終的に 2009 年 10 月に NGS 側からセンサス局に対して提供された^[注 13] [(50) p.6]。

6. 住所の点検作業 (address canvassing)

2010 年センサスの最終準備工程として位置づけられているのが、あらゆる既知の街路や構築物等を確認し、新たな道路等を点検し、人々の現住地あるいは居住の可能性のある構築物を特定し、センサス局が保有する住所リストと地図の更新を行う住所点検 address canvassing (address verification と呼ばれる) である。それは、住戸や集団施設に対してセンサス調査票を正確に郵送あるいは調査員が手渡すのに必要な郵送リストや所在地の確認を目的として行われたものである。センサス実施の際には、この点検作業によって最終的に確認された住所に調査票が郵送され、一部の非都市地域については、調査員が訪問による調査票の手渡しを行った。

センサス局は、12 の常設の地域センサスセンターの他に、2010 年センサス実施業務の遂行のために 450 の出張所を地方出先機関として設置した。2007 年夏に行った GPS 取得試験結果や 2008 年秋から 2009 年初頭にかけて実施された検証テストを経て、151 の出張所の管内においてセンサ

スの調査時点(2010年4月1日)の約1年前である2009年3月30日から7月中旬にかけて住所の点検作業が実施された[(28) p.2、(46) p.2]。

住所の点検作業には、リスター(lister)と呼ばれる約14万人の現地作業員が、センサス局臨時職員を示す白いセンサスバッジを付け黒いセンサスカバンを持って従事した。なお、リスターの中には、ACSの調査員も含まれている。

彼らはGPS機能(ArcPad)を装備した[図5]のようなモバイル端末を携帯して、合計1億4500万の住戸や集団施設の住所、所在地の確認を行った。なお、このモバイル端末は、海外のメーカーによって特注生産されたもので、特に高齢のリスターでも操作が容易なように設計されていた。

端末に表示される地図に載っていない道路や施設をリスターが住所点検作業の過程で新たに発見した場合、端末を用いて街路をトレースするとともに、通りに沿った住戸の正確な位置情報の収集が行なわれた。また、高層ビル等の構築物の陰など立地上の理由でうまくGPS座標情報が取得できなかった住戸については、取得できた近隣の位置情報を用いて補完処理が行われた。このようにして取得された地図の更新情報や位置情報がTEDに格納された[(37) p.5]。

この住所確認作業でのリスターの最大の任務は、域内の各住戸のGPS情報を取得することであった。リスターは各住戸の正面玄関から40フィート以内で情報端末の地図画面をチェックすることで、その取得を行うように予め指示されていた[(34) p.4]。実際のaddress canvassingにおいては、特に高齢のリスターの中には事前に指定されていた現場に立つことなく車の中あるいは歩道からGPS座標情報の取得を行う者もいた。さらには、域内の事情に詳しいと称するリスターの中には、現地に赴くことなく地図に直接入力を行うケースもあったと報告されている。しかし、センサス局での事後評価では、大半の作業については満足すべき精度でGPS情報の取得は行われたとのことである[(50)p.5]。

なお、このような作業に係る収集情報の秘密保護並びに作為による誤情報入力を防止するために、各リスターには、端末のログオンに必要なパスワードと親指の指紋による生体認証ソフトが組み込まれた情報端末が配布された[(46) p.1]。

リスターはまた、端末経由で閲覧可能な調査区の住戸のMAF住所を確認するとともに、その住人が当該住戸の他に住宅を保有しているかどうかについてのヒアリングを行った[(34) p.6]。これは、把握漏れあるいは重複把握防止対策の一つとして行われたものである。アメリカには、例えば、北部諸州の在住者で冬季間フロリダ州など温暖な地域で過ごすsnowbirdsと呼ばれている人口集団が存在する。このような人口を「6ヶ月以上の居住地」において確実に把握するために、彼らの主たる居住地を確認するのがその主たる目的であった。

7. 2010年センサスの実施状況

2010年センサスでは、2009年度内に既に開設済みの150か所に加えて新たに開設された344か所の地域センサス事務所が調査地域実施本部となった。約1億2200万世帯に対しては郵便による送付・回収方式で、950万世帯に対しては調査員による調査票の手渡しによる留置き調査方式で、さらに約140万の世帯ならびに施設等では調査員による実査が実施された。その後、特に回答率の低い地区に対しては4月中旬にかけて調査票の再送も行われた。

センサス調査票の送付を受けた世帯の約35%は所定の期限までに回答を提出しないであろうと事前に予想^[注14]されていたが、郵送による最終的な回収率は、ほぼ前回のセンサス並みの74%に

達した〔(48) p.11〕。約 4700 万世帯に上る未回答世帯に対しては、約 60 万人の調査員が直接訪問による督促調査(nonresponse follow-up(NRFU))を実施した〔(48) p.1〕。

以前のセンサスにおいて調査員が調査督促のために訪問した世帯の中にはすでに回答を済ませていた世帯が少なからずあった。調査員が調査区を訪問する前に未回答の世帯を事前に特定することができれば、このような調査の非効率あるいは重複調査の可能性を回避することができる。このため、2010 年センサスでは、郵送されるセンサスの調査票に予めその対象世帯の所在位置情報を示すバーコードがプリントされた。回収済み調査票のバーコード情報を用いて照合することで、TIGER 地図上に未回答の世帯の位置情報を表示することができる〔注 15〕。報告期限の終了後から 7 月 10 日にかけて、全米で 58 万人の調査員が、調査票にプリントされた地点バーコードとの照合により判明した未回答世帯が居住すると考えられる住戸を正確に特定し、ピンポイント訪問による回答の督促が行われた〔注 16〕。

むすび

本稿では、アメリカの 2010 年センサスの実施計画を事例に、世帯が居住する住戸等の GPS コード情報が調査の回収精度の向上並びに調査実施経費の削減のために活用されていることを見てきた。

住戸等の GPS 座標情報を予め取得しておくことで、それを用いて統計単位の把握率(回収率)の改善を実現することが期待できる。それは、調査員の目視に依存する調査ではしばしば見過ごされがちであった統計単位の発見という把握漏れ対策、回答済み単位に対する重複調査の回避といった調査実務の効率化という点で人口センサス以外の例えば企業調査の精度向上にも活用できる可能性を持っている。

フランス国立統計経済研究所(INSEE)人口社会統計局人口部がその整備、更新を所管し、住戸だけでなく事業用建物もカバーした建物登録簿(répertoire d'immeubles localisés: RIL)は、GPS コードをすでに保有している〔(2)〕。また、オランダ中央統計局(CBS)の企業統計部門でも、統計の作成面で GPS コードの多角的な活用が試みられている。

すでに〔森(3)〕でも論じたように、GPS コードを付与することで、既存の個体レコードは、それまで利用上の制約となっていたものを解除し新たな統計情報の利用価値を拓くだけでなく、統計の作成面でも有効性を持っているように思われる。この分野における海外統計の新たな展開方向の中には、将来わが国においても採用が検討されてしかるべきものも少なくないように思われる。新たな統計の展開方向の一つとして、今後もその動向に注視したい。

〔注〕

1)総務省統計局では、2011 年 3 月 11 日の東日本大震災による津波の罹災世帯・人口規模を、水没地区の航空写真画像から取得した地図情報とその前年 10 月に実施された国勢調査の基本単位区人口データを同定することによって推計し、公表している。なお、URL 中の番号 02(青森県)、03(岩手県)、04(宮城県)、07(福島県)、08(茨城県)、12(千葉県)に対応している。
<http://www.stat.go.jp/info/shinsai/pdf/sinsui02.pdf>(2011 年 8 月 15 日現在)

2)例えばホワイトハウスは GPS によれば緯度(38.898590)、経度(77.035971)で示される。これは、

分以下の緯度・経度情報をラジアンに換算表示したものである。GPS 情報は無理数であり、特定の桁数で表示される情報は厳密にはポイントを示すものではなく微小な面を表すラスタ型情報である。

3) Waite, Preston J. and Birnbaum, Nicholas I., (2003) “Census 2000 Methods and the Vision for the 2010 Census”, Workshops on Population Census and Census Micro Data. Japan Statistics Research Institute, Hosei University

4) アメリカの 2010 年人口センサス実施の予算総額は約 115 億ドルであった。仮に従来のような詳細調査併用型の調査方式を 2010 年調査でも採用した場合、約 14 億ドルの予算が新たに必要になるといわれている〔(16)〕。

5) 不変価格で比較した 1 世帯あたりの調査票回収費用は、1990 年センサスでは 0.88 ドル、2000 年調査では 3.19 ドルであった。

6) 2010 年センサスの準備過程で寮、グループホーム、受刑施設、ホームレスシェルター等の集団施設 (group quarters) について所在地の住所確認作業が行われた〔(46) p.1〕。

7) ZCTA は米国郵政公社の配達住所地域を表す USPS ZIP コードと同一のものではなく、それに近似させることを意図して設定された準統計地域区分である〔(47) p.2〕。

8) 「(住所や特定の位置を含め) 個人や事業所についての情報は、統計的集計値としての提供以外の形では提供することはできない。」〔(35) p.2〕

9) 2000 年センサスの実施の際には 1600 万枚の地図が作成された〔(37) p.34〕。

10) 「CE95 7.6m」とは、政府が設定している地図の精度規準 (National Map Accuracy Standard: NMAS) のレベルの一つで、地図上のポイントの 95% が実際の位置から誤差 7.6m の範囲内に収まる精度の地図を意味する〔(37) p.26〕。なお、センサス局地理部によれば、このレベルでの精度設定は、概ね道路幅を想定したものとのことである。

11) Baker 社が受注したプロジェクトには、他に以下のような企業が参加した。Dynamix Corporation, SMW Engineering Group, Inc., International CADD Service Inc., Woolpert LLP, RETTEW Associates Inc., Photo Science Inc., Satellite Data Systems Ltd., Applied Geo Technologies Inc.

12) Booz, Allen, Hamilton 社

13) 事後補正作業は次の 3 つの過程で行われた。

① 取得データの ORACLE からソートした ASCII ファイルへの変換

② NGS による補正プログラムを用いた ASCII ファイルの処理

③ NGS から提供された補正データの ORACLE への再変換

なお、全体で 6000 万以上の GPS 座標情報の補正が行われた〔(50)p.6〕

14) 例えば、フロリダ州 Brevard 郡の 2010 年 4 月 27 日現在の回収率は 77% (2000 年調査の回収率は 70%) であった〔(32) p.1〕。

15) 2004 年にニューヨーク市とジョージア州で実施された非回答のフォローアップ (NRFU) 試験調査の一環として、調査員に該当する住戸の地点情報を地図上に表示するために GPS 座標が収集された〔(50)p.4〕。

16) 政府責任局 (GAO) は、NRFU が、回答内容の品質よりも作業の迅速さに重きが置かれていたとして、2020 年センサスに向けてその改善の必要性を提言している〔(49)〕。

付図

(1) 現実と乖離した TIGER マップの例



緑のラインは旧 TIGER による街路。住宅の列と対応しておらず、一部の道路は水路(図中の黒い部分)の上にはみ出している。

出所[(11) p.14]

(2) 現実と乖離した TIGER マップの例

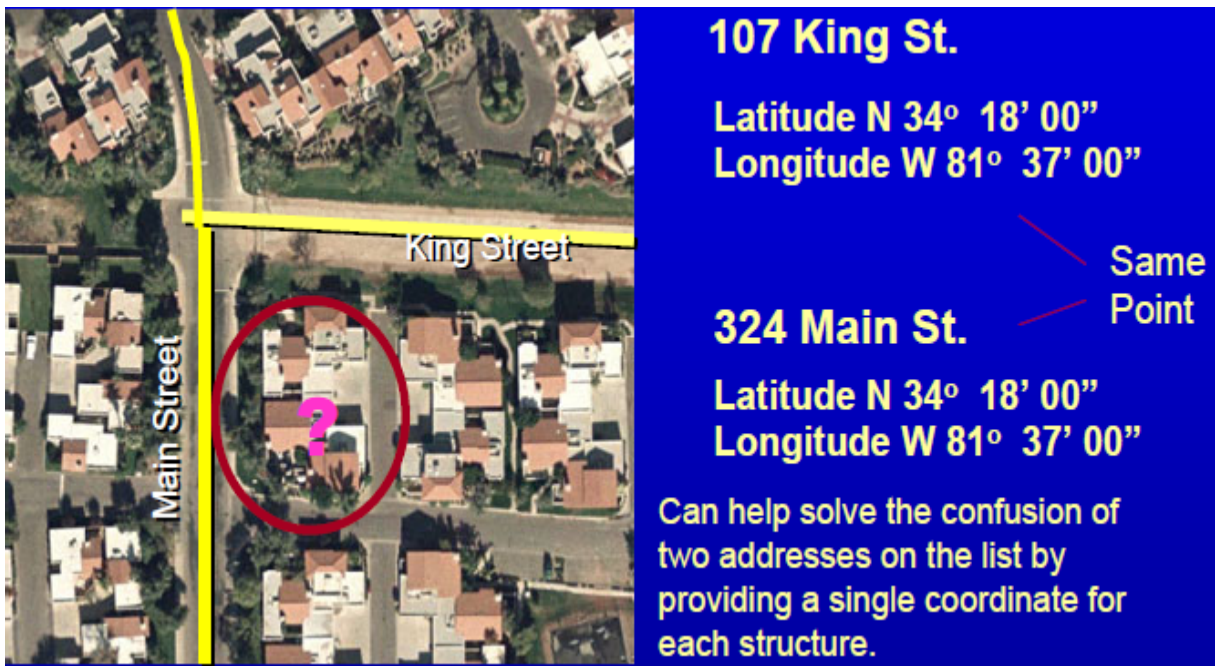


赤いラインで表示されているのが旧 TIGER による街路、白いラインが実際の街路※図の左の矢印部分では 10m、また右の矢印部分では 100m ものズレ

出所[(11) p.9]

なお、原資料 p.9 の図の脚注では赤と白による道路表記に関するコメントが誤記されている。

(3) GPS 情報による MAF 住所の二重登録の検出



107 King St.
Latitude N 34° 18' 00"
Longitude W 81° 37' 00"

324 Main St.
Latitude N 34° 18' 00"
Longitude W 81° 37' 00"

Same Point

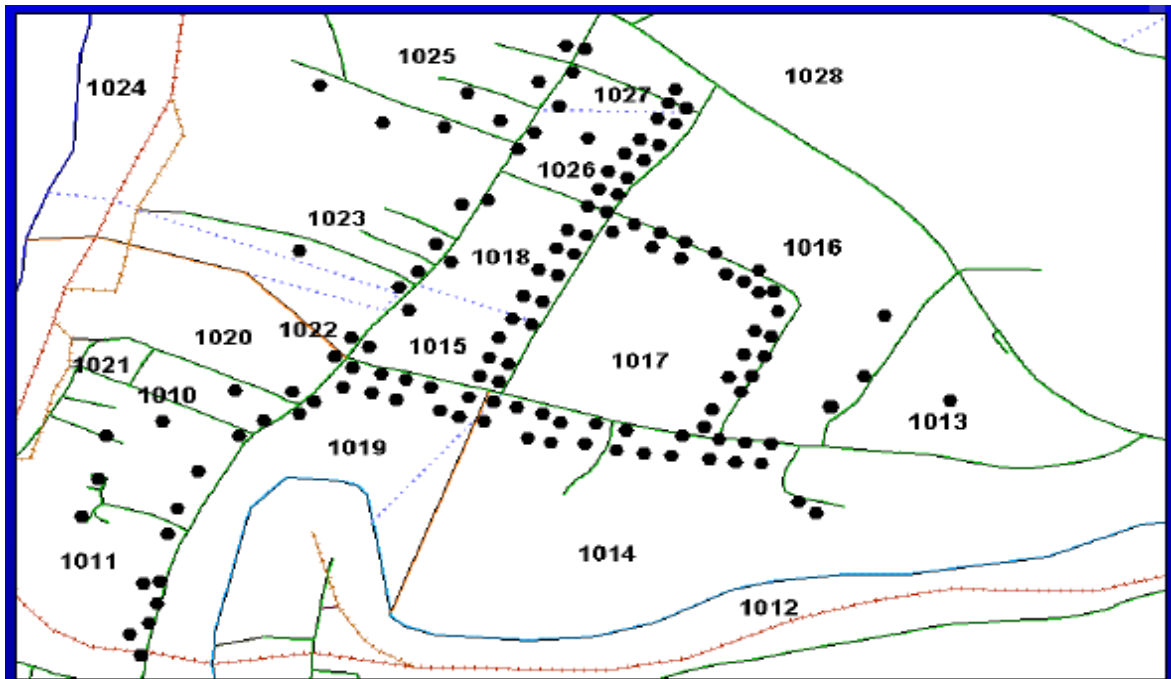
Can help solve the confusion of two addresses on the list by providing a single coordinate for each structure.

同じ緯度経度情報(緯度:北緯 34 度 18 分 00 秒、西経 81 度 37 分 00 秒)を持つ住戸が MAF に 2 つの異なる住所として登録されている。その結果、調査票が同一の住戸に重配される可能性がある。

出所 [(11) p.15]

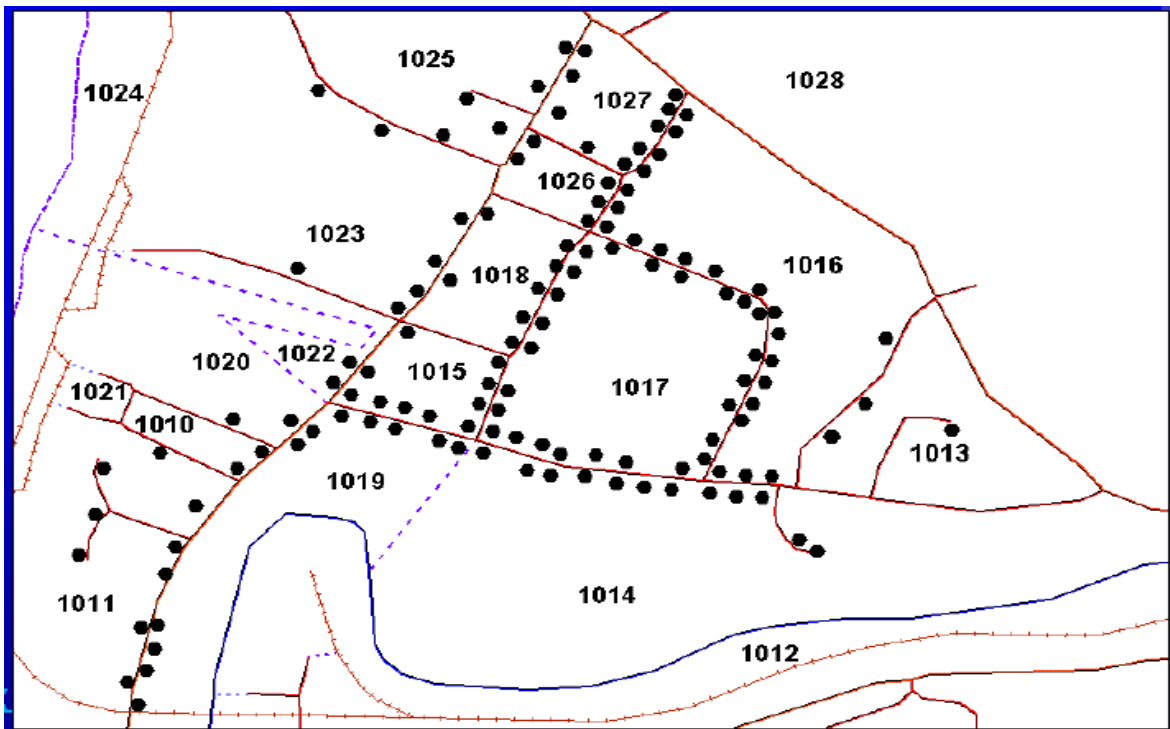
(4) センターライン情報の精度向上による TIGER 地図の改善

[改善前の TIGER マップ]



出所 [(11) p.16]

[改善後のマップ]

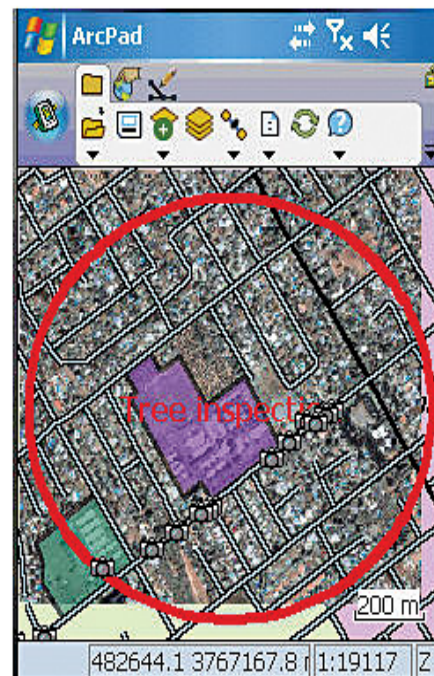


出所 [(11) p.21]

(5) ArcPad 機能付端末



With ArcPad 8, users can now synchronize field edits directly with the enterprise geodatabase via ArcGIS Server.



Identify target areas for inspections and GIS data collection.

(出所) [(25) p.1]

【文献】

- (1)森博美(2009)「統計個票情報の情報特性について」『経済志林』法政大学経済学会 第76巻 第4号
- (2)西村善博「フランスの新人口センサスにおける建物登録簿(RIL)の意義と課題」『経済論集』熊本学園大学 第15巻 第3・4合併号 2009年3月
- (3)森 博美(2011)「統計調査における地点情報の把握による統計の情報価値の新たな展開可能性について」『経済志林』法政大学経済学部学会、第78巻第3号
- (4)Robert W. Marx (2000) “A Vision for a 21st Century MAF/TIGER”, <http://www.census.gov/Geo/mod/vision.pdf>
- (5)U.S. General Accounting Office(2002) *2000 Census – Best Practices and Lessons Learned for More Cost-Effective Nonresponse Follow-up* (GAO-02-196)
- (6)U.S. Department of Commerce Office of Inspector General(2003), “MAF/TIGER Redesign Project Needs Management Improvements to Meet Its Decennial Goals and Cost Objective” (Public Release) Final Report OSE-15725 September 2003
- (7) Waite, Preston J. and Birnbaum, Nicholas I.,(2003) “Census 2000 Methods and the Vision for the 2010 Census”, Workshops on Population Census and Census Micro Data. Japan Statistics Research Institute, Hosei University
- (8)U.S.Census Bureau (2004) “MAF/TIGER Redesign – Project Overview”. http://www.census.gov/geo/mtep_obj2/obj2_issuepaper12_2004.html (2010年4月現在)
- (9)Michael Baker Corporation(2004), “Baker Awarded GPS Data Collection Contract from U.S. Census Bureau’s Geography Division”. http://www.mbakercorp.com/index.php?option=com_Content&task=view&id=1891&temid=203 (2010年4月現在)
- (10)All Business com(2004) “Census Bureau eyes GPS for 2010 U.S. effort”. (2010年4月現在)
- (11)Fusaro, Randy (2004) “Spatial Accuracy Matters: TIGER Accuracy Improvement”. http://www.eurostat.net/km_pub/no48/workshops_docs/positional_accuracy_dit_march_05/S1_Key_Note_Address/by-Randy-Fusaro.pdf
- (12)Wendy Hawley (2004) U.S. Census Bureau’s MAF/TIGER Accuracy Improvement Project – Objectives, power point presented at WaTrans Partner’s Meeting on March 2nd.
- (13)Timothy Trainor (2004),”Update on the Census Bureau’s MAF/TIGER Enhancement Program (MTEP)”,<http://training.esri.com/campus/library/Bibliography/RecordDetail.cfm?ID=35939>
- (14)Wendy Hawley (2004) “MAF/TIGER Accuracy Improvement Project– An Update”, presented at WaTrans Partner’s Meeting on March 2.
- (15)Preston Jay Waite and Burton H. Reist (2005) “Reengineering the census of population and housing in the united states”, *Statistical Journal of the United Nations Economic Commission for Europe*, Vol.22 No.1
- (16)U.S. Government Accountability Office (2006) “2010 Census: Census Bureau Needs to Take Prompt Actions to Resolve long-standing and Emerging Address and Mapping Challenges”. <http://www.gao.gov/products/GAO-06-272> (2010年4月現在)
- (17)Charles Louis Kincannon (2007),”Agency Response to Assistant Inspector General for

Systems Evaluation”.

(18)U.S. Census Bureau (May 2,2007), “Locally Hired Census Workers to Update Address List, Visit Every Residence in San Joaquin County”, http://www.census.gov/Press-Release/www/Releases/archives/2010_census/010026.html

(19)U.S. Census Bureau (2007), *Census 2010:Key Challenges to Enumerating American Indian Reservations Unresolved by 2006 Census Test*, final report No.OSE-18027.

(20)Philip Lutz (2007), “2010 Census and Local Partnerships”, <http://www.docstoc.com/docs/5393600/2010-CENSUS-AND--LOCAL-PARTNERSHIPS/>

(21)U.S. Census Bureau (2007) “The 2010 Census” prepared statement by Charles Louis Kincannon before the Subcommittee on Federal Financial Management, Government Information, Federal Services, and International Security, U.S. Senate.

(22)U.S. Government Accountability Office(2008), *2010 Census – Census Bureau’s Decision to Continue with Handheld Computers for Address Canvassing Makes Planning and Testing Critical*.

(23)Campaign For Liberty (2009), “2010 census using GPS to pinpoint addresses”, <http://www/campaignforliberty.com/blog.php?view=15607>

(24)Eric Zuehlke (2009),”What You Need to Know About the 2010 U.S. Census”, <http://www.Prb.org/Articles/2009/2010censuspolicyseminar.aspx>

(25)ArcNews Online (summer 2009) ”ArcPad 8 Now Available–Enhances Use Experience and Improves Customization”, <http://www.esri.com/news/arcnews/summer09articles/arcpad8.html>

(26)ArcNews Online (summer 2009) ”ArcPad Eases Address Canvassing in 2010 U.S. Census”, <http://www.esri.com/news/arcnews/summer09articles/arcpad-eases.html>

(27)JETL(July 22 2009), “Big Brother’s Countdown: Census 2010 via GPS”, <http://jetl.wordpress.com/2009/07/22/big-brothers-countdown-census-2010-via-gps/>

(28)U.S.Census Bureau News (March 25, 2009),”Census Workers to Verify Addresses in the U.S. First Major 2010 Census Operation to Employ 140,000”, http://www.census.gov/Press-Release/www/releases/archives/2010_census/013441.html

(29)Timothy F. Trainor (2009),”Census Geographic Programs Update”, presented for the National States Geographic Information Council 2009 Midyear Conference on February 24”.

(30)G.O’Hara, Charles and Cheriadat, Anil (2009) “Accuracy Analyst: A Horizontal Accuracy Assessment Tool” *Accuracy Analyst White Paper*

(31)Elaine Griffin (2010),”U.S. Census 2010 Overview – Major Changes to the 2010 Census and How It Affects You” <http://americanaffairs.suite101.com/article.cfm/census-2010-overview>

(32)Brevard county (April 27,2010),“Brevard County’s Census Awareness Campaign”, <http://Brevardcounty.us/census/>

(33)Minnesota State Council on Disability (May 1, 2010),”2010 Census – Be Counted!”, <http://State.mn.us/portal/mn/jsp/content.do?contentid=53619252&contenttype=EDITORIAL&hpage=true&agency=MSCOD>

(34)WorldNet Daily (May 2, 2010), “Census GPS-tagging your home’s front door”, <http://www.Wnd.com/?pageID=97208>

(35)U.S. Census Bureau “Background Information on the MAF/TIGER Accuracy Improvement

Project (MTAIP) and Government Partnership”. <http://www.census.gov/geo/mod/background.html> (2010年4月現在)

(36)MAF/TIGER Modernization – CBD Announcement of a Request for Information (RFI)

(37)Thu Nguyen, “Project 3 TIGER Modernization Project”. <http://test.scripts.psu.edu/users/t/ttn143/Project3.html>

(38)“ArcPad Eases Address Canvassing in 2010 U.S. Census” <http://www.esri.com/news/arcnews/summer09articles/arcpad-eases.html>

(39)Census Bureau, “Census Bureau launches 2010 Census Advertising Campaign”, <http://www.prnewswire.com/news-releases/census-bureau-launches-2010-census-advertising-campaign-81490202.html>

(40)U.S.Census Bureau, “GPS address location”, <http://2010.census.gov/mediacenter/2010-census-timeline/address-canvassing/gps-address-location.php>

(41)U.S. Census Bureau, “Common Privacy Questions – Be assured, your information is safe”, <http://2010.census.gov/2010census/privacy/common-privacy-questions.php>

(42)Decennial Census Progress Review and Program Replan

(43) U.S. Census Bureau, “2010 Decennial Census Local Update of Census Addresses (LUCA) Program”, <http://www.lpgov.org/Census/LUCA%20Info.htm>

(44)Census Bureau, “2010 Census Progress”, <http://2010.census.gov/2010census/how/2010-census-progress.php>

(45)Census Bureau, “Take 10 Map”, <http://2010.census.gov/2010census/take10map>

(46)U.S.Department of Commerce, “U.S.Census Bureau News”, <http://2010.census.gov/news/press-kits/one-year-out/address-canvassing/address-canvassing-facts-statistics.html>

(47)Wikipedia, “Topologically Integrated Geographic Encoding and Referencing”, http://en.wikipedia.org/wiki/Topologically_Integrated_Geographic_Encoding_and_Referencing

(48)US Census Bureau,(2010) *U.S. Census Bureau FY 2010 Performance and Accountability Report*

(49)U.S. Government Accountability Office (2010) “2010 Census: data Collection Operations Were Generally Completed as Planned, but Long-standing Challenges Suggest need for Fundamental Reforms”, <http://www.gao.gov/products/GAO-11-193>

50)Aaron Jensen,(2010) “Current Status of the Global Positioning System Program at the Census Bureau Geographic Division”, discussion paper.

[謝辞]

本論文は、平成 22 年度日本学術振興会科学研究費補助金 基盤 B「政府統計データのアーカイビングシステムの構造と機能に関する国際比較研究」(課題番号 22330070 研究代表者森博美)による研究成果の一部である。

なお、本稿の一部は、2011 年 8 月 11 日にアメリカ商務省センサス局地理部 (Geography Division)において菅幹雄教授(法政大学経済学部)、宮川幸三准教授(慶応義塾大学産業研究所)とともに行ったインタビュー調査結果に依拠している。業務多忙な中、長時間にわたるわれわ

れの調査に快くご協力いただいた Timothy F. Trainor (Division chief)、Michael Ratcliffe (Assistant division chief)の両氏をはじめ、Geography Division の多くの職員の方々にこの場を借りてお礼を申し上げたい。

オケージョナル・ペーパー(既刊一覧)

号	タイトル	刊行年月
5	人口動態統計における交通事故死亡統計の特徴について	2000.05
6	Trends in U.S. Working Hours since the 1970s	2001.07
7	わが国における外国人の国籍別出生率について	2001.09
8	東京の消費構造－東京都生計分析調査	2002.10
9	Wide Variations in Statistics Data Sets on the Same Subjects－Reconsidering the Report of the Indian National Statistical Commission	2003.12
10	日中 1995 年産業別購買力平価の推計	2004.04
11	日本における「統計法」の成立	2005.06
12	「統計法」と法の目的	2005.07
13	諸外国におけるマイクロデータ関連法規の整備状況とデータ提供の現状	2005.09
14	統計に係る個人情報の秘密保護について	2006.08
15	若年層における雇用状況と就業形態の動向－『就業構造基本調査』のマイクロデータによる実証分析	2006.12
16	社会生活行動から見た若年層の不安定就業化・無業化の分析	2008.03
17	国勢調査による従業地把握の展開と従業地別就業データの意義	2009.06
18	無償労働の評価と世帯生産サテライト勘定	2009.10
19	エンゲルとザクセン王国統計	2009.12
20	第一次統計基本計画と政府統計の直面する課題	2010.01
21	エンゲルとプロイセン統計改革	2010.02
22	エンゲルと 1875 年ドイツ帝国営業調査	2010.03
23	調査形態論再論	2011.03
24	統計を規定する諸要因との関連から見た時空間個体データベースの可能性について	2011.04
25	位置情報を用いた調査票情報の情報価値の拡張とその分析的意義について	2011.06

オケージョナル・ペーパー No.26

2011年9月15日

発行所 法政大学日本統計研究所
〒194-0298 東京都町田市相原 4342
Tel 042-783-2325、2326
Fax 042-783-2332
jsri@adm.hosei.ac.jp
発行人 森 博美