

オランダの virtual census について

森 博 美

要 約

1960年代末以降、欧州各国を中心にプライバシー意識が高揚する中で、オランダは1971年センサスを最後に調査員による人口把握(enumeration)の中止に追い込まれる。同国では1981年、1991年センサスについては、人口登録情報と労働力調査などいくつかの標本調査データに基づきセンサス結果表の作表が行われた。センサスのための特別な調査が実施されないことから、それらは virtual census と呼ばれてきた。ただし、これらの virtual census については、人口登録を母集団情報として標本調査の結果を一定の乗率を用いて復元することで必要なセンサス項目についての集計を行っただけのもので、その結果精度については、政府や研究者からも重大な疑問が投げかけられていた。そこでオランダ統計局(Statistics Netherlands)は、2001年センサスに向けてレジスター情報と標本調査から得られる情報を個体レコードベースで照合・連結し、さらに項目分類等に関する調整、補正を施すことで、センサスデータの精度向上のための方法論の開発を行った。これによって、同国の virtual census は新たな段階へと入ることになる。本稿では、virtual census がどのような登録情報や調査データからいかにしてセンサスで要請されている結果表が作表されているかという virtual census の作業過程を点検することを通して、その意義あるいは方法論上の問題点を明らかにした。

現在、世界の人口センサスは、一方では、わが国のような調査員による直接的な人口把握という伝統的なセンサス並びに北欧諸国に見られるレジスターベースの人口把握を両極とする主軸を持ち、他方ではイギリスやフランスに見られるような標本調査の活用形態を副次的軸に持つ多様な展開をみせている。このような多様なセンサスの展開が究極的にある特定のセンサス方式へと収斂するかどうかについて、現時点でそれを見通せるような段階にはない。しかし、virtual census の方法論は、人口規模さらには利用可能な行政情報の存在といった一定の条件の下で、それまでのセンサスの諸形態とは異なる新たなタイプのセンサス実施の可能性を示唆しているように思われる。それは、現在、世界的規模で進行中のセンサスの展開の空白を埋めるとともに、センサスの将来の展開方向を模索する上で有効な情報の一つとなりうるものと期待される。

はじめに

1990年以降、オランダでは政府業務に係る行政負担の効率化が推し進められた。オランダ統計局の統計作成に係る行政事務量は1994年から2002年の9年間に6割削減され [(11) p.4]、また2002年から2006年の間にさらに25%の削減 [(11) p.3] が政府の政策目標として掲げられた。その過程で、同局は、一貫して統計予算、人員の削減を余儀なくされてきた。

オランダ統計局は、一方で増大する回答拒否に象徴される深刻化する統計調査環境、他方で統計作成コストの削減と報告負担の最小化という条件の下で、多様でしかも詳細かつ相互に整合的な統計情報の提供を要請されてきた。このような新たな統計作成の局面のなかで同局は、伝統的な統計調査のみに依拠した統計作成方式だけではそれへの対応はもはや困難であるとして、1990年代初頭に行政情報の統計作成への活用を開始した。その後、情報技術の進歩を背景に、使用可能な行政情報の範囲を大幅に拡充することで、統計作成業務の本質的な転換をはかった。同局では、関係法規を抜本的に改定し、統計以外の目的のために収集、作成されている行政情報の統計作成への積極的利用 (large-scale statistical recycling of information) を基盤とする統計政策へと転換した。その結果、2001年に同局は、統計調査によるデータの収集を、調査以外の手段では入手できない情報のみに限定するという調査実施の新たな原則を打ち立てた。これによって同局は、それまで調査統計に置いていた統計作成の軸足を行政情報へと改めることになった。このようなオランダにおける統計政策の転換をまさに象徴するのが、1981年以来導入されてきた virtual census の方法論の抜本的再構成 (reengineering) である。

本稿では、オランダ統計局が2000年ラウンド世界人口センサスとして2001年に実施した virtual census の方法論を検討するとともに、その特徴並びに問題点を明らかにしてみたい。

1. 伝統的センサス方式から virtual census へ

オランダでは、1829年に第1回のセンサスが実施されている。初期のセンサスの実施を所管していたのは内務省で、同省はその後1989年まで10年周期で合計七度のセンサスを実施している。1899年にオランダ統計局が設置され、その後センサス実施業務は同局へと移管される。オランダ統計局は、1899年、1909年、20年、30年、47年、60年、71年、81年、91年、2001年と十度のセンサスを実施してきた。このうち1971年までのセンサスは、伝統的な調査員による実査 (enumeration) による調査であった。

同国におけるセンサス実施の方法転換の最初の転機となったのは、1971年センサスである。

60年代末以降のプライバシー保護意識の高まりの中で、センサス情報が各自治体で管理している人口登録 (Population Registers: PR) の登録内容の点検と更新に使用されている点が問題視された。政府業務の電算化が進展する中、George Orwell 流に「Big Brother がわれわれを監視している」[(1) p. 2] という形で反センサス運動¹⁾が展開され、また、政府が行政情報として把握している項目を何故センサスとして改めて調査する必要があるのかといった点でもセンサスは批判的の的となった。結局、同国の当時の人口の約 2.3%にあたる 29.6 万人が 1971 年センサスで把握漏れとなったが、そのうちの 2.3 万人は回答拒否によるものであった。こういった回答拒否者に対して罰則を適用するかどうかは政府内で検討されたが、罰則を科すのは非現実的であるということで結局その適用は見送られた [(3) p. 4]。

1970 年センサス法は、1971 年センサスの次のセンサスの実施を 1981 年としていた。しかし、1979 年に二度にわたって実施されたセンサスの試験調査では、回答拒否率が全国平均で 26%²⁾に達し、また、センサス態度調査でもセンサスへの協力度が著しく低位であることが確認された。さらに 1980 年には、膨大な調査経費をかけてこのようなセンサスを実施するのが果たして合理的なのかどうか、また調査の実施から結果の公表までに要する時間の長さといった伝統的な実査方式でのセンサス実施の問題点などを取り上げたモノグラフも公刊された。このような一連の動きの中でオランダ中央統計局は、伝統的な実査方式での 1981 年センサスの実施を最終的に断念した [(3) p. 4]。

一方で統計調査への国民の協力度が急激に低下し、他方で調査実施に対する予算面での厳しい制約が課せられる中で、Eurostat 等からのセンサス結果表の提供要請への対応を迫られたオランダ統計局は、1981 年と 1991 年センサスについては、PR の登録情報と労働力調査 (Labour Force Survey: LFS)³⁾ データから必要なセンサス結果表⁴⁾の作表を行った。

しかし、これらのセンサスで作表に当たったの原情報として使用された PR と LFS それぞれのデータは相互に有機的に統一されたものではなく、単に PR からは得られないセンサス項目について、LFS データを PR を母集団情報として復元推計しただけのものに過ぎなかった。そのため、得られた推計値の精度については、研究者だけでなく政府関係者からも疑問視されるものであった。さらに、市レベルでの地域集計表も作成されず、自治体での利用ニーズに全

1) 第二次世界大戦中にナチスが人口センサスの個票情報をユダヤ人の居住地の特定に使用したこともセンサス反対運動の一因であるといわれている。

2) 試験調査での拒否率は地方部での最低 21%から最高は都市部での 37%の範囲で分布した。

3) 1981 年と 1991 年の労働力調査の抽出率はそれぞれ、5%と 1%である。

4) 性、年齢、世帯属性等のセンサス変数については PR 情報が、また就業状態や教育に関する結果表の作表には専ら労働力調査データが用いられた。

く応えられないという決定的な欠陥を持っていた。

オランダにおけるセンサスの第二の転機は、1991年に訪れる。1980年代にオランダでは、伝統的なセンサスに復帰するかあるいはミニセンサスと呼ばれる標本調査のいずれの方式を採用するかをめぐっての議論が行われた。しかし、結果的にはいずれの方式も不採用となり、1991年には同国のセンサス法が廃止されることになる。オランダ統計局は、センサス法の廃止と引き換えに広範な行政記録の作成権限を獲得することになる⁵⁾。

2. virtual census の新展開 2001年センサス

1981年と1991年にいずれも virtual census として実施されたオランダのセンサスは、2001年センサスにおいて新たな展開を見せる。そこでは、個々の推計結果の間に整合性を持たせるために、各種の登録情報と標本調査データが個体ベースでリンクされ、各変数間の定義の統一や欠損値や欠損データの補定といった micro-integration と呼ばれる一連の編集作業が行われた。なお、この作業工程そのものは、virtual census の精度改善を一つの目的にオランダ統計局が新たに構築することになった社会統計データベース (Social Statistical Database: SSD) の整備作業の一環として行われたものでもあった。なお、SSD の構造や機能については、機会を改めて詳細に論じることにしたい。

3. virtual census の情報源情報

EU 諸国については、Eurostat が 2000 年ラウンドセンサスの実施について、*Statistical Requirements Compendium* の中で、各国に報告を求める統計の表章単位、把握時点、製表リスト等に関するガイドラインを設定している。その中で Eurostat は各国に対して、合計 40 の結果表の所定期限内の報告を求めている。その内訳は、職業や学歴、経済活動といった人口関係の表が 30 表、住宅関係 (8 表)、通勤関係 (2 表) である。このようなセンサス結果表の提出要請は、2000 年ラウンドセンサスについては、あくまでもボランティアな性格のものとし

5) 2003年11月に裁可され、同年12月18日に公布された84条からなる改正統計法 (Statistics Netherlands Act) では、第5章(第33条~第52条)がデータの獲得、使用並びに提供にあてられている。特に第33条は、統計局長に政府等の公的機関が業務遂行に関して維持している登録データの統計目的での使用、請求権限を、また第34条課税・社会保障番号の登録への入力及び統計目的での使用権限を付与している。

オランダの virtual census について

て行われたもの [(8) p.3] であるが、オランダ統計局では、それに対応することを前提に 2001 年 virtual census センサス計画が立案された。

2001 年センサスの集計表の作表に使用されたデータの多くは、オランダ統計局が構築、維持管理しているから得られる。しかし、SSD が持つ全ての変数 (統計項目) が作表に使用されているわけではない。ちなみにオランダでは、Eurostat が各国に提出を要請した合計 40 の表のうち、14 の表 (No.1, 3, 4, 5, 12, 13, 17, 18, 20, 23, 30, 33, 35, 40) については、PR など既存の登録データから作成でき、住宅関係の 8 つの表 (No.14, 24, 25, 26, 27, 28, 37, 39) については住宅登録と住宅調査データから、そして残る 18 の表については、SSD, LFS, それに (Survey on Employment and Earnings: SEE) から得られるデータに基づき推計、作表された [(12) p.5]。

2001 年センサスでは、SSD に対する情報の供給源ともなっているファイルも含め、合計 7 種類の登録情報と 2 種類の標本調査データがセンサスの情報源情報として使用された。

(1) センサスの基幹情報としての人口登録 (PR)

オランダにおける 2001 年 virtual census の基盤情報となっているのが PR である。オランダでは、1994 年 10 月 1 日に包括的な人口登録システムとして (Gemeentelijke Basis Administratie persoonsgegevens: GBA) が、社会保障登録、水道・電力供給登録、外国人に対する警察の登録、それに高齢者年金基金登録を統合した人口登録 (PR) として整備された。なお、オランダでは PR についての地方分散型での維持、更新の原則が立法化されていることから、中央人口登録にあたるものは制度的には存在しない [(12) p.1]。

PR がセンサスに代わって同国における最も包括的でしかも有効な人口把握情報とされるのは、次のような事情による。すなわち、一方で PR から得られる人口数は、行政面では、各自治体に対して国から交付される補助金配分の基礎数字として用いられる。他方、国民の側でも、選挙権だけでなく、失業保険や医療、さらには児童扶助や子供の就学といった様々な政府サービスの受給申請のためには、人口登録番号の記載が求められることになっている。また、外国人についても、合法的な在留のためには、地域の所轄警察への届出による登録が義務づけられている。このように、行政と住民いずれの側にとっても PR は、登録行為並びに登録ファイルの管理のためのインセンティブが作用している。不法在留外国人など登録から漏れている人口も一部存在するが、その規模は全人口の 1%未満であると見積もられている。

(2) その他の登録情報

2001年センサスでは、PRの他にも雇用並びに社会保障に関する各種の行政登録情報が使用された。雇用保険登録（雇用者）（Employee Insurance Scheme Registration System-Employees: EIS-Employees）、雇用保険登録（失業保険）（Employee Insurance Scheme Registration System-Unemployment Insurance: EIS-UI）、雇用保険登録（障害保険）（Employee Insurance Scheme Registration System-Disablement Insurance: EIS-DI）、自営利益最終所得税評価登録（Register of final income tax assessments on profits of self-employed persons: FiTap）、社会保障給付（Social Assistance Benefits Administration: SABA）、それに所得税情報を持つ財政データベース（Fiscal Database: FiBase）の中の租税事前支払登録の6種類の登録情報がそれである。

オランダでは、1990年代に雇用保険制度に基づき、雇用者についての包括的な雇用保険登録が（Employee Insurance Scheme Registration System: EISRS）として構築された。この雇用保険登録は、年齢、性、人種、出生地、現住所、婚姻、さらには就業といった情報を持っている。なお、EISRSは就労時間と就労場所に関する情報を持たないことが、この登録情報をセンサスのための情報源として使用する際には一定の調整が必要となる。

また、FiTap登録ファイルからは、自営業従事者の就労日数と産業部門に関する情報が得られるが、この登録については二つの問題点が指摘されている。第一は、FiTapが所得発生期間に関する正確な情報を持たないことである。このため、2000年のいずれかの期間に自営活動の結果何らかの所得が発生した者については、センサス基準時点（2001年1月1日）現在での自営業従事者とみなされる。この「みなし」によって、結果的にセンサスでは自営業従事者数が過大評価されている。もう一つの問題は、課税額に関して税務当局との間での見解の相違から税額が未確定であることからFiTap登録から漏れている自営業主の存在である。後者については、2001年センサスにおいて約4万人（自営業主全体の5%に相当）が把握漏れとなっていると見積もられている。

SABA登録ファイルは、社会保障の種類、支給金額、支給期間といった情報を持ち、FiBase登録ファイルは、税金の支払いと関連した就労と社会保障データ、さらには年金や生命保険情報も保有している。これらのファイルは、例えば、センサス把握時点で55-64歳の非労働力人口は、FiBaseデータによって年金のみを所得の源泉とする者と生命保険受給者として求められるなど、年金受給者、退職者など非労働力人口の把握に使用される。なお、FiBaseのデータは、micro-integrationと呼ばれるマッチング済みレコードの各変数値の補正の過程で就業や給付に関する欠損データの補正のための補助情報としても使用されている [(7) p. 247]。

オランダの virtual census について

ちなみにオランダでは、個人ベースの登録情報にはすべて SoFi が付与されており、SoFi をマッチングキーとして相互にマッチングできるようになっている。なお、これらセンサスにおいて使用されている登録情報のうち EIS-UI, EIS-DI, SABA についてはリンクされたレコードについて就業と給付情報に基づき必要なセンサス項目に関するデータの推計のための補助情報として専ら使用されるものであり、センサスの結果表の作表には直接使用されてはいない [(7) p. 248]。

(3) 調査データ

センサスに使用されている上記の各種登録からは、個人の学歴や職業それに勤め先の規模、給与、所定内労働時間、就労の場所といった就業に関する情報、さらには住宅の室数、所有形態、建築年に関する情報が得られない。このために 2001 年センサスでは、上記の登録情報と併せて、標本調査として定期的に行われている労働力調査 (Labour Force Survey: LFS)、雇用・所得調査 (Survey on Employment and Earnings: SEE)、居住環境調査 (Survey on Housing Conditions: SHC) から得られる調査データがセンサス結果表の作表に用いられた。

まず、職業と学歴情報については、オランダ統計局はそれらに関するデータを LFS から入手している。なお、この他にも LFS は、失業者あるいは通学や家事に専従する非労働力人口の把握にも用いられている [(12) p. 3]。一方、勤め先の規模、給与、所定内労働時間、就労の場所といった就業に関する情報については、SEE から得られるデータが使用されている。さらに、センサスの結果表のうち住居建物の建築年、住宅の所有形態、住居の室数といった住宅関連の表については、2001 年 1 月 1 日現在の PR と 2001 年の住宅登録それに 2000 年の SHC から得られるデータが使用された [(10) p. 5]。

これらの調査のうち LFS は、標本数約 10 万の標本調査として実施されているものである。この程度の標本規模ではセンサス結果表としての必要な推計精度の確保が困難であることから、2001 年センサスでは 2000 年調査 (標本数約 12 万) と 2001 年調査 (標本数約 11 万) という 2 年分のデータをプールすることで使用された。また、個人の就業情報を提供する SEE は、各企業の賃金台帳に基づく調査である。大規模企業については全数、一方、中小企業については標本調査となっている [(7) p. 246]。

(4) センサス使用情報源一覧

表 1 は、2001 年センサスの主要情報源として使用された登録、調査データについて、その種類別にレコード数やセンサス項目等を示したものである。

表 1 2001 年センサスで使用した情報源一覧

情 報 源	単 位	基準時点における 記録総数・標本数	基準時点	センサス項目
1. 人口登録	人	1600万レコード (690万世帯レコード)	2001年1月1日	性 年齢 出生国 国籍 1年前の居住国 居住地域 続柄 婚姻状態 家族の状況 核家族世帯 一般世帯の構成 世帯の状態 世帯規模 子供の数 世帯以外の居住者数
2. 雇用保険登録 - 雇業者 (EIS-Employees)	職	650万レコード	2000年12月 22-31日	雇業者 産業分野 (NACE) 社会保険料控除前賃 金総額 (補助変数)
3. 雇用・所得調査 (SEE)	職	標本 300万レコード	2000年12月 22-31日	雇業者 普段の労働時間 就業の場所
4. 自営利益最終所得税評価 登録 (FiTap)	自営業者	79万レコード (把握漏れ約5%)	2000年額	自営業従事者 産業分野 (NACE)
5. 雇用保険登録 - 失業保険 (EIS-UI)	給付	44万レコード	2000年額	非センサス項目
6. 雇用保険登録 - 障害保険 (EIS-DI)	給付	100万レコード	2000年額	非センサス項目
7. 社会保障給付 (SABA)	給付	58万レコード	2000年額	非センサス項目
8. FiBase 登録 (租税事前支払登録)	所得移転	720万レコード (雇 用者) 270万レコード (年 金、生命保険)	2001年12月31 日	雇業者か否か 年金 / 生命保険給付 (退職者)
9. 労働力調査	人	標本 12万レコード (2000年)、11万レ コード (2001年)	2000年と 2001 年	学歴 職業 失業者 就学者 家事従事者

[出所] Eric Schulte Nordholt et. al. (2004), *The Dutch Virtual Census of 2001*, p. 247

4. 基盤情報と各登録・調査データのリンケージ

PR が把握した人口をベースに、それに登録情報と調査データをオランダの社会保障番号 (Social security and Fiscal number: SoFi)⁶⁾ をキー変数としてマッチングし相互にリンクすることでセンサス結果表の作表のための原データ (個体レコード) ファイルが作成される。一方、SoFi 番号が付与されていない標本調査の個人レコードの場合、SoFi による照合ができないことから、性、出生年月日、住所 (4 桁の数字と 2 桁の文字からなる住所コード) 情報をマッチングキーとして PR とレコードのリンケージが行われている [(7) p. 248]。

マッチングによる個体ベースでのレコードのリンケージ、さらには micro-integration と呼ばれているリンク済みデータの定義調整や欠損データ等の補定は、SSD のデータ更新の中心的作業に属する。ここでは一部のセンサス結果表の作表に SSD の個体レコードが用いられている点だけを確認し、リンケージと micro-integration の方法の詳細については、機会を改めて論じることとしたい。

5. センサス変数の定義調整

Virtual census のバックボーン情報である PR にリンクされた各種の登録情報と調査データからなるデータベースについては、その情報源が多様であることから、定義や概念に関する調整が必要となる。Micro-integration と呼ばれるこの作業工程の詳細については別稿に譲ることにして、以下ではセンサスデータの作成に関する限りでのデータ処理内容の概要について紹介しておこう。

センサスで求められる個人情報については、PR とそれにリンクさせた調査データあるいは行政情報から得られる。しかし、いくつかの表では世帯に関する表章が必要となる。PR では、登録者全員に個人番号 (personal identification number: PIN) が与えられており、各個人レコード (personal list: PL) には、同居の有無を問わず親、子供、配偶者についての PIN 情報が与えられている。従って、PIN 情報を手がかりに PR から世帯等を再構成することができる [(12) p. 3]。

世帯の構成に際しては、最初に住所情報を基に施設等の居住者に関するレコード群を特定し、残りのレコード群から PIN を手がかりに一般世帯 (private households) が構成される。同一

6) SoFi は 9 桁の数字のみから構成される。なお、現在は、civil service number と呼ばれている。

住所を持つレコードが全て PIN によって家族と同定された場合それが単一の家族とされ、関連付けられないレコードの者については、その住居の同居人として扱われる。親と子からなる核家族が世帯のベースとされ、親の兄弟姉妹あるいは祖父母等はその世帯に同居する家族員となる。二世帯同居世帯については、若い方の夫婦を中心とする核家族に両親が家族員としての同居者として扱われる。また、兄弟姉妹のみからなる住居についても、親の PIN 情報を通じて家族関係が特定され、一つの世帯として扱われる。例えば、両親とその子供、祖父母及び 2 人の同居人からなる住居については、夫婦・子供・2 名の家族員からなる世帯と 2 つの単独世帯として扱われる [(12) p. 4]。

PR の各個人レコードである PL が有する続柄 PIN 情報を用いて一般世帯の約 93% の世帯を確定できる [(7) pp. 245, 255]。しかし建物やアパートを共同使用している家族以外の同居人のいる住戸については、居住者の関係の特定は容易でない。法的な婚姻関係にない子供無しの男女同居者などを PR 上で特定するのが最も困難であるといわれている。このためオランダ統計局では、住戸を 6 に類型化⁷⁾し、PL 情報を用いて対数回帰モデルによる世帯類型の推定が行われている。なお、現地調査での筆者の質問に対してインフォーマントからは、この処理作業は一応は自動化されているものの、事柄の性格上そのアルゴリズムは複雑なものとならざるを得ず、なお改善の余地があるとの回答であった。

就業については有業者、無業者別の表章が行われるが、有業者の中には同時に複数の仕事に従事している者もいる。副業を持つ有業者については、最大の収入を得ている仕事をその者の主業とみなしそれ以外を副業とするという原則が採用されている。なお、産業分類や職業分類上の格付けは主業に従って行う [(7) p. 255]。

2001 年 1 月 1 日をセンサスの把握時点⁸⁾とすることから、同日現在のデータがセンサスの基盤情報として PR から取り出される。他方、就業データについては、12 月 22 日が基準日 (reference day) として選ばれる。これは、多くの企業がクリスマス休暇に入る 12 月の最終週から 1 月 1 日にかけての就業状態は日常とは異なり、また期間契約の就業者の中には雇用契約が年末以前に終了するケースが少なくない点等を考慮したものである [(10) p. 6]。なお、LFS

7) 同一の住戸に 2 人が同居、同一住戸に 3 人が同居、同一住戸に 4~9 人が同居、同一住戸に片親家族と 1 名が同居、同一住戸に人組の男女と 1 名が同居、上記 ~ のうち複数の郵便配達単位を持つ住戸 [(12) p. 4]

8) オランダでは、毎年 2 月 15 日現在で把握した登録人口を各自治体の公式の人口としている。2001 年センサスでは 2002 年 1 月 1 日までに利用可能な全ての情報を結果数字に反映させる措置がとられた。このために、センサスの人口数は 2001 年の公式登録人口と全国で約 1,500 人異なるといわれている [(12) p. 2]。

オランダの virtual census について

で失業者とされた者について、2001年1月1日現在でのSSDのレコードが異なっている場合には、後者の情報が優先される [(8) p.5]。また、経済活動についてオランダ統計局では、次のような処理のルールも定めている。

0 - 3 歳の者 (その他の経済的非活動)、4 - 15 歳の者 (有業者である場合にもフルタイム就学者)、75 歳以上の者 (退職者扱い)、65 - 74 歳の無業者 (退職者)、55 - 64 歳で専ら年金受給者又は生命保険受給者 (退職者)、16 - 74 歳で週1時間以上就労する者 (経済活動人口 (雇用者))、16 - 74 歳で自営の者 (経済活動人口 (自営)) [(7) pp.256-7]

なお、SSD にレコードを持つ 16 - 64 歳の有業者でも退職者でもない者が LFS において失業中と回答している場合、失業給付受給の有無にかかわらずその者は失業者として扱われる [(7) p.257]。

就業者の勤務先住所地については、次のようなルールでその確定が行われる。まず、雇用主の場合、本所 (本社事業所) の所在地が勤務先住所地となる。一方、雇用者については、その者が複数の事業所を持つ企業に勤務する場合、個々の事業所の雇用者数を考慮しつつ本人の現住所から直線距離で最短の事業所の所在地がその者の勤務先住所地となる。なお、当該事業所における勤務者数が雇用者数を上回る場合、各事業所の雇用者数と整合的であるように二番目に近い事業所がその者の勤務先となる [(12) p.5]。

表 2 人口関係の変数一覧

個 人				世 態			
変 数 名	単位	レベル数	源 泉	変 数 名	単位	レベル数	源 泉
性別	人	1	PR	世帯主	世帯	2	PR
年齢	人	7	PR	世帯の種類	世帯	3	PR
世帯特性 (*)	人	3	PR	家族の子供数	世帯	6	PR
世帯の種類	人	5	PR	25 歳未満の子供の数	世帯	6	PR
世帯の規模	人	3	PR	18 歳未満の子供の数	世帯	6	PR
1 年前の居住地	人	2	PR	6 歳未満の子供の数	世帯	6	PR
国籍	人	5	PR	世帯の種類と配偶者			
出生国	人	1	PR	の経済状態	世帯	3	LFS + SSD
地域 (NUTS)	人	5	PR	経済活動者の数	人	2	LFS + SSD
学歴	人	5	LFS + PR	世帯の規模	人	数量	PR
就業状態	人	3	LFS + SSD	世帯の数	世帯	数量	PR
従業上の地位	人	7	LFS + SSD				
職業	人	4	LFS + SSD				
就業形態	人	3	SEE + SSD				
産業 (NACE)	人	6	SSD				
人数	人	数量	PR				

表注 (*) 離別、死別等の世帯特性

[出所] Pieter Everaers and Paul van der Laan (2000), The Virtual Census, p.266 より作成。

既に述べたように、LFS はオランダの virtual census において重要な情報源の一つとなっている。しかし、LFS をセンサスの情報源として用いるに当たって大きな制約がある。それは、LFS で施設居住者が調査の対象外されていることである。その結果、施設居住者については、センサス項目である、学歴、就学、家事、それに職業や失業に関する情報が全く得ることができない。このためオランダ統計局では、施設居住者が一般世帯員と同一の分布構成に従うとの強い仮定をおいて、これらの項目に関する欠損データの補定を行っている [(12) p.5]。

前頁の表 2 は、人口センサスで作成が Eurostat から求められている人口関連の表 (30 表) の製表に用いられている変数、変数区分レベル数、およびデータの源泉を示したものである。

6. センサス結果表の作成 (1) データブロックと結果表

センサスの対象となる人口は、雇用者や雇用主、自営業主といった経済活動人口と退職者や通学・家事専業従事者といった非労働力人口、さらには 15 歳以下の人口からなる。こういった人口の各セグメントについて、いろいろなセンサス項目について集計表が作表される。すでに述べたように、集計表の中には、PR の情報だけから作表できるものもあれば、他の登録情報あるいは調査データに基づいて作表されるものもある。次のブロック 1～ブロック 6 は、集計に用いられる情報源とそこから積み上げ (推計) られるセンサス項目の関係を示したものである。

[ブロック 1] 登録ブロックで、PR と SSD から得られる全ての変数を含む。この中には当該年齢階級の総人口 12,036,171 人が含まれる。このデータブロックに属するデータだけから作成できる集計表は、単に計数を積み上げることによって作成される。

[ブロック 2] NACE (産業部門) ブロックとも呼ばれるこのブロックは、非雇用者並びに NACE が判明している雇用者・雇用主に関するデータから構成される。このブロックデータは、登録情報だけから得られる変数を産業部門分類と組み合わせることで作成できる表の作表に使用される。

[ブロック 3] このブロックは EcAct (Economic Active people) ブロックと呼ばれ、雇用者、雇用主、退職者及び LFS から得られるこれらのいずれにも該当しない個人をカバーしている。このブロックデータは、登録情報だけから得られる変数と従業上の地位との組み合わせによって作成できる表の作表に使用される。

[ブロック 4] 第 4 ブロックは SEE ブロックと呼ばれ、調査から得られるデータの中では最大のデータブロックを構成する。それは、SEE から得られる 15 - 74 歳の雇用者並びに退職者

及びその他の者を合計したものである。なお、SEE データブロックからは従業員規模に関する情報が得られ、従業員規模と登録が有する情報だけから作成できる表の作表に使用される [(2) p. 268]。

[ブロック 5] このデータブロックは SEE/NACE ブロックと呼ばれ、NACE ブロックの中で特に SEE の情報を有するレコード群から構成される。このブロックは、産業部門並びに従業員規模と登録情報のみから作成される表の作表に使用される。

[ブロック 6] 第 6 ブロックは LFS ブロックと呼ばれる。それは、LFS から得られる全変数と登録情報とから作成される学歴や就業形態、職業等の集計事項を持つ表の作表に用いられる [(2) p. 269]。

図 1 は、15～74歳 人口について、集計結果表の使用変数と集計に使用されるデータブロックの対応関係を示したものである。

図 1 集計表と対象データブロック (15 - 74 歳人口)

	ブロック1	ブロック2	ブロック3	ブロック4	ブロック5	ブロック6
	登録ブロック	NACEブロック	EcActブロック	SEEブロック	SEE/NACE ブロック	LFSブロック
雇用者		NACE		SEE	SEE	LFS
雇用主		NACE				LFS
退職者		NACE 非適用		従業員規模 非適用	従業員規模 とNACE 非適用	
その他			LFS			LFS
	12,036,171	11,985,413	8,793,530	8,144,956	8,109,261	163,741

[出所] Pieter Everaers Paul van der Laan (2000), The Virtual Census, p. 268 より作成。

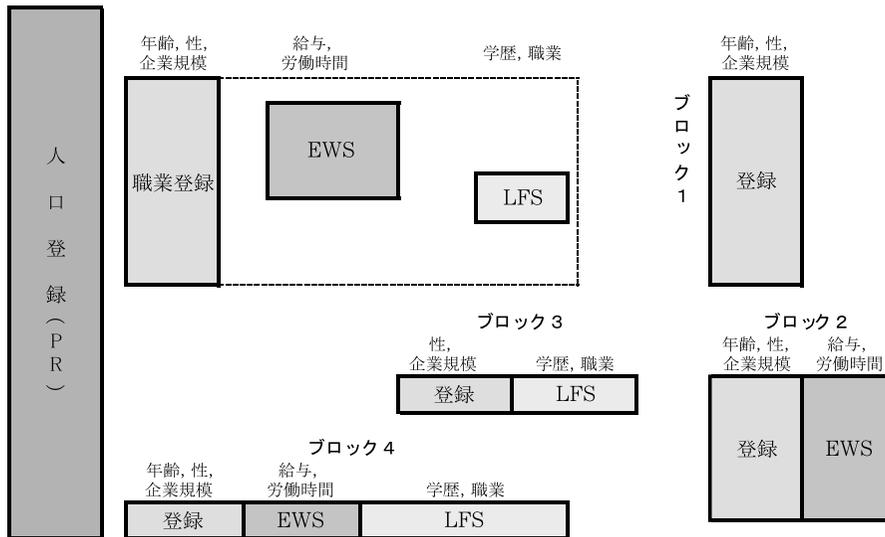
7. センサス結果表の作成 (2)

SSD は、雇用者についての就業ファイルと雇用主及び雇用者のいない自営業主ファイルとをデータベースの中に持っている。これらの人口規模の特定は次のように行われる。

まず雇用者の就業ファイルでは、EIS-Employees から得られる就業者数、SEE と FiBase

登録情報をもとに雇用者数が推計される。このうち SEE は所得と雇用者の就業に関する情報を持つことから、センサス項目の中の所定内労働時間と就業の場所のデータがこれから得られる。なお、SEE は、大規模企業については全数で、小企業については標本調査として実施されている。一方 FiBase 登録からは就労所得と社会保障所得が得られ退職者数の推計に使用されているほか、小企業従事者についての欠損データの補定にも用いられている [(12) p. 2]。

図 2 登録情報と調査データのリンク例



[出所] M. Houbiers (2002), Towards a social statistical database and unified estimates at Statistiek Netherlands, p.3 より作成 (一部加筆)

PR に他の登録情報や調査データをリンクすることで構築されるデータベースでは、センサス結果表の作表に必要な項目 (変数) の組を持つレコード群が大小様々な方形 (rectangular) を形成するが、作表にあたっては、それらの中で利用可能な最大の方形が使用される。ちなみに、この方形は、登録だけから構成されるものと登録にリンクされた一つ以上の調査データからなるものがある。このうち登録だけから作表できるものについては登録から直接導出されるが、学歴や職業など標本調査データを用いるものについては推計作業が必要となる。

推計に当たっては元データにウエイトをかけて母集団への復元がはかられる。例えば、性・年齢・学歴別の多元集計表のように、専ら LFS 結果データを用いて作表される場合、非回答分を考慮して算定される publication weights を乗率として母集団サイズへの復元が行われる。なお、このウエイトは initial weights と呼ばれており、標本調査データについてはこの

publication weights が initial weights となる。LFS については、安定的な推定結果を得るのに必要な標本数を確保するために 2000 年と 2001 年の調査データをプールして使用していることから、センサス日の総人口の約 2 倍となるようにウエイトは設定されている⁹⁾。なお、登録データについてはその倍率は 1¹⁰⁾である。

登録と調査データから構成されるデータブロックから作表する場合、あるいは年齢のように同一の変数が複数の種類の異なる年齢階級区分を持つ位階的なカテゴリー区分をとりうる変数を含む表、さらには多次元のクロス表を相互に整合的なものとして作表する場合には、さらにこれらの調整のために追加的なウエイトの付与が必要となる。例えば、LFS が与える学歴・国籍別人口の数値は、登録における国籍別人口とは必ずしも一致しない。また、性・年齢(各歳)・学歴(12 区分)・国籍(3 区分)の集計表は LFS からしか作表できないが、性・年齢・国籍別人口に関する集計表は、登録からも得られる。これらの表が相互に整合性を持つためには、前者の表のうち学歴を除く周辺数について、登録から得られる後者の各セルの値と整合的である必要がある [(2) p.272]。virtual census においてこのような作表過程での推計結果の積み上げの整合性をはかるために用いられているのが、反復加重法 repeated weighting (RW) という補正手法である。

RW は、次のように使用される。まずマッチングされた登録と調査データより構成されるレコード群から作表に必要な変数を持つ最大の方角がデータブロックとして特定される。次いで、この特定されたデータブロックが母集団の規模を反映するようにウエイト (initial weight) が与えられる。その後、単純集計表からクロス表、さらにより高次元のクロス表へと

9) LFS に initial weights として適用されるのは、

N : 2001. 1. 1 人口

n_{00} : 2000 年の LFS のサンプル数 (正確には回答数)

n_{01} : 2001 年 "

$w_{00,i}$ ($i=1 \cdots n_{00}$) 各 LFS の publication weight

$w_{01,i}$ ($i=1 \cdots n_{01}$)

このとき、2000 年の LFS のウエイトには

$$\frac{N}{\sum_{i=1}^{n_{00}} w_{00,i}} \times \frac{n_{00}}{n_{00} + n_{01}} \text{ が、また 2001 年の LFS には } \frac{N}{\sum_{i=1}^{n_{01}} w_{01,i}} \times \frac{n_{01}}{n_{00} + n_{01}}$$

がそれぞれウエイトとしてかけられる。それぞれ第 1 項は 2001 年 1 月 1 日の人口に一致するような調整項で、第 2 項は 00 と 01 年のウエイトを平均化する項である。例えば、15~74 歳については、2000 年の LFS 結果については 0.56 倍、2001 年のそれについては 0.52 倍されねばならない [(2) p.270]。

10) 登録の把握時点がセンサス日の 2001 年 1 月 1 日と異なる場合は、総数を同日の登録数に総数を一致させるように調整される。

作表工程が積み上げられるが、その際には、より低次元のクロス表の各セル値が次の次元のより高次元クロス表における周辺数と一致するよう継続的な補正が加えられる。

以上のような作表過程は自動化されており、オランダ統計局が独自に開発した (Vullen Reference Database: VRD) というソフトを用いて行われる。作表は、1つの世帯表と4つの個人表 (15歳未満人口, 15~74歳施設人口を含む総人口, 15~74歳一般世帯人口, 75歳以上人口) の合計5つのデータブロック別に、それぞれVRDを用いて行われる。なお、VRDが使用できるためには、各データブロックでは、作表に用いられる各変数について、全てのカテゴリー (複数の異なるレベルのカテゴリー区分をもつ場合にはそれらの全てのカテゴリー) が特定され、また複数の異なるレベルをもつ変数については、レベル間の階層性が確定していなければならない。

virtual census の作表過程では広範にRWが適用されているが、RWの適用の有無による推計結果の差異は必ずしも大きくはない。initial weight だけによる推計結果とRWを追加的に施したそれとの差異は、ほとんどの場合、1%程度であるとされている¹¹⁾ [(2) p. 274]。

ところで、データブロックから目標とするセンサス結果表まで積み上げ作表する過程では、いくつかの問題が生じうる。

第一は、いわゆる sampling zero の問題である。もともと母集団において稀少なケースについては、標本調査データから作表された表のセルの数値がゼロとなることが少なくない。その場合、いかなるセルもゼロ以外の数値となるように変数を減じより低次元のクロス表として表示するか、あるいはある変数のカテゴリーの一部を統合することで、このようなゼロの発現を防止する措置が取られている [(2) p. 271]。

11) initial weights repeated weights による推計結果の比較

推計値の単位は1,000人

変数及びカテゴリー	男			女		
	SW 推計値	RW 推計値	$\frac{RW-SW}{SW}$ (%)	SW 推計値	RW 推計値	$\frac{RW-SW}{SW}$ (%)
ISCO - 1 議員、上級行政官、管理的職業従事者	699.2	695.6	- 0.5	229.4	231.1	0.7
ISCO - 2 専門的職業従事者	710.7	705.4	- 0.7	493.5	499.8	1.3
ISCO - 3 技術者及び準専門的職業従事者	613.0	607.8	- 0.8	633.1	640.9	1.2
ISCO - 4 事務従事者	272.4	271.9	- 0.2	569.4	569.4	0.0
ISCO - 5 サービス職業従事者、店舗等販売業従事者	257.6	259.2	0.6	546.7	541.5	- 1.0
ISCO - 6 熟練農林漁業職業従事者	77.7	78.3	0.8	27.4	27.0	- 1.5
ISCO - 7 熟練職業従事者	678.5	677.3	- 0.2	34.7	34.8	0.3
ISCO - 8 装置機械操作員、組立工	400.7	398.8	- 0.5	47.6	47.9	0.6
ISCO - 9 初級・単純職業従事者	266.0	272.4	2.4	253.2	250.5	- 1.1
ISCO - 0 軍隊	34.4	34.2	- 0.6	2.8	2.8	0.0
分類不能の職業	277.7	287.2	3.4	269.1	261.2	- 2.9
非就労	1,759.5	1,759.5	0.0	2,881.9	2,881.9	0.0

[出所] Pieter Everaers and Paul van der Laan (2000), The Virtual Census, p. 274 より作成。

また、例えば LFS データからは「71 歳」の「他の EU 国籍」を持つ「男」についての推計値は得られない。このような標本調査データから直接推計できないセルについては、性・年齢・国籍、年齢・学歴・国籍、性・年齢・学歴、性・年齢(5 歳階級)・学歴・国籍という 4 つの表を使ってその人数の推計が行われる [(2) p.272]。

第二は、変数間の相互連関の問題である。例えば、従業上の地位と就業状態の各カテゴリーは相互に独立ではなく、経済活動(3 区分)による失業者数は従業上の地位(4 区分)での失業者数と正確に一致しなければならない。このような変数間の関係については、VRD ソフトはサポートしていない。このため、変数間の依存関係に起因する非整合的な要素については、自動化された作表結果に手作業による追加的な補正が加えられる。

この他にも、調査あるいは登録における項目区分がセンサスの結果表として Eurostat から提出が求められているそれと整合的でない場合がある。そのために全てのセンサス項目の区分を全て網羅的に表章できない場合、オランダ統計局ではその旨の注記を行うことで対処している [(12) p.5]。

8. センサスの地域表章と結果表の表示形式

EU 諸国については、Eurostat が 2000 年ラウンドセンサスの実施について、*Statistical Requirements Compendium* の中で、各国に報告を求める統計の表章単位、把握時点、製表リスト等に関するガイドラインを設定している。このためオランダ統計局では、PR の登録情報を広範に活用しているほか、人口・世帯属性並びにその経済活動等に関する個体データベースとしてすでに稼働中の SSD が保有するデータを Eurostat のガイドラインの要件に適合する形に調整することで必要な作表を行っている。

Eurostat が 2000 年ラウンドセンサスにおいて EU 加盟各国に対して提出を要請している 40 の結果表について、地域表章レベル別にその内訳を見ると、28 が全国表で、残りの 12 表は地域別表となっている。また、地域別集計表の内訳は、Eurostat が定める地域表章分類体系 NUTS¹²⁾ に従って NUTS3 に該当する同国の地域研究プログラム調整委員会 COROP¹³⁾ によ

12) NUTS とは、Eurostat が定める地域区分 (Nomenclature d'Unités Territoriales Statistiques: Nomenclature of Territorial Units for Statistics) で、各国について位階的な地域区分のレベルが定められている。(http://en.wikipedia.org/wiki/Nomenclature_of_Territorial_Units_for_Statistics 参照) なお、NUTS は、各国における行政区分とは必ずしも一致しない場合がある。ちなみに、オランダの NUTS 区分レベルと該当する地域数は次の通りである。

る地域区分である COROP region レベルの結果表 (9 表), NUTS5 レベルの表 (3 表) となっている。また, オランダ統計局では, 2000 年ラウンドセンサスとして Eurostat から報告が求められているこれらの表に加えて, 第二次都市評価指標プログラム (Urban Audit II)¹⁴⁾ に従って Eurostat から提供を求められた国内十大都市¹⁵⁾ については, 住宅及び労働に関する区レベルの小地域表章を持つ集計表も作成している [(10) p.4]。

集計表の中には, 多くの集計項目を持つ多重クロス表あるいは地域別表の場合, 集計対象数に比べて相対的にセルの数が多いため, 結果的に稀少なケース数を持つセルが発生する可能性がある。このようなセルについては秘密保護の見地からも問題があるとして, virtual census では, 稀少ケースのセルを含む結果表の表示については次のような原則を定めている。

まず, 結果表の全てのセルでケースが 25 を超えている場合, その表は無条件で公表される。次にケース数が 10 未満のセルが含まれる結果表の場合, 隣接する他のセルと統合して表示される。さらにこれらの中間にあたるケース数が 10~24 のセルが含まれる結果表については, 表に含まれるセルの半数以上が 25 以上のケース数を持ち, しかも 10 未満のセルが出現していない場合に限り, ケース数 10~24 についても公表扱いとなる [(2) p.275]。

NUTS 区分	NUTS1	NUTS2	NUTS3	NUTS4	NUTS5
地域の名称	地域ブロック Landsdelen	州 Provincies	COROP Regions	-	自治体 Municipalities
地域数	4	12	40	-	450

オランダでは全土が東部, 西部, 南部, 北部の 4 つに区分 (NUTS1) されており, NUTS3 として編成された COROP regions には 10 前後の自治体が含まれる。

なお, 国によっては NUTS 区分の中で該当するものが存在しない場合がある。小国の場合, 上位の NUTS が全て 1 つの地域として表章される。ちなみにオランダでは, NUTS4 は存在しない。

- 13) Coördinatie Commissie Regionaal Onderzoeks Programma (Coordination Commission Regional Research Programme) が Eurostat の地域区分基準などを受けてオランダ国内の主としてオランダ統計局によって分析目的のために使用される地域区分を定めるもので, 各地域レベルが COROP 地域と呼称される。
- 14) 都市評価指標 (Urban Audit) は, 欧州の主要都市における生活の質 (QOL) の評価についての要望の高まりを受けて欧州委員会 (EC) の地域政策総局 (DG REGIO) と Eurostat が策定したプログラムで, EU 加盟諸国並びに加盟候補国の都市圏についての信頼できる比較可能な情報の提供を目的としている。1990 年代半ばのパイロットフェイズ (Urban Audit I) では 58 都市が, また第二フェイズ (Urban Audit II) では 258 都市の 2001 年データが整備された。
- 15) Eurostat では, 大都市統計 (Urban Audit) として, 域内の大都市についての統計情報の提供を求めている。オランダでは, Amsterdam, Rotterdam, The Hague, Utrecht, Eindhoven, Tilburg, Groningen, Enschede, Arnhem それに Heerlen が指定大都市として指定されている [(7) p.10]。

9. 2010 年ラウンドセンサスに向けて

Eurostat は EU 加盟各国政府に対して 2010 年ラウンドセンサスでは 125 表の提出を提案しており、2008 年 9 月を目途とした決着を目指して、目下各国の統計当局と Eurostat との間で折衝が進行中である。このような折衝過程を経て最終的に報告書が取りまとめられ、各国に結果表の提出が要請されることになる。このような中、8 月末時点でオランダ統計局では、Eurostat からの要請に対してどの範囲での対応が可能であるかについて、財務当局との間で最終的な詰め協議が進行中とのことであった。なお、前回のセンサスで提出が求められた 40 表はあくまでも Eurostat 側からの希望として提起されたものであったのに対し、2010 年ラウンドセンサスでは、それらの表の提出が義務づけられる点が新たな展開として注目される。これを受けてオランダ統計局では、結果表提出の義務化を背景に、財政当局に対して予算の拡充に向けた要請が行われている。

このような Eurostat 側から提起されている新たな要求に対して、オランダ統計局内部ではそれへの対処策が具体的に検討され始めている。新たに提出が要請されている集計表のいくつかについては、RW の手直し等も含め、現行の SSD からの作成可能性が具体的に検証され、すでに対応への一定のめどがついている。その一方で新たに提案されている 125 表のうちどの範囲まで対応可能であるかについては、なお不確定要素が残されている。最終的にどれだけの種類の結果表の提出が求められるかについては現時点では確定していないものの、2001 年 virtual census に比べ数倍の結果表の提出が要請されることは間違いなく、作表作業に投入すべき人的、財政的資源の拡大は避けられないことから、統計局では提出の「義務的性格 (mandatory)」を根拠に財政当局に対してそれに向けての予算措置を強く要請している。

オランダでは、2010 年ラウンドセンサスについても、virtual census 方式を採用する方針は確認されており、2011 年に予定されているセンサスでも、把握すべき項目 (変数) そのものはほぼ前回センサス並みとされている。なお、2001 年センサスの新たな展開としては、オランダ統計局が地理登録情報 (Land Register data) の活用を計画している点が注目される。これは、住戸や事業所について、それまでの所在地情報に加え GPS 等の地理空間情報を活用することで、レコードのマッチングや世帯の再構成の精度向上だけでなく、勤務先データの改善などを通じて virtual census の結果精度の改善に向けた取り組みと考えられる。

10. virtual census の意義と問題点

Virtual census の場合、センサス項目に関する情報収集のための特別な調査の実施あるいは行政情報の整備を必要としない。このため、報告者だけでなく他の行政当局に対して追加的な負担を強いることなく Eurostat が求める結果表を作成することができる。このような報告負担や実査に基づくセンサスの企画・実施に伴う業務負担の回避もさることながら、virtual census 導入の最大のメリットは財政面での経費削減効果にあるとされている。財低的な慧眼もちなみに、伝統的な実査に基づくセンサスの場合には3億ユーロ相当の経費が必要とされるのに対し、virtual census では300万ユーロとその作成に関わる経費を1/100にまで削減できた [(10) p. 4]。

2001年センサスはvirtual censusとして伝統的な実査によるセンサスに対して経費等の削減効果の面で評価されているだけでなく、virtual census それ自体としても、それまでの1981年、1991年センサスに対して推計結果精度の面で質的な改善が図られた [(7) p. 258]。結果精度の改善は、次のような制度面並びに作表方法や推計方法の改善によるものである。まず、制度面での改善としては、1996年に改正統計法施行され、オランダ統計局は政府が保有する行政情報に対するより広範なアクセス権が制度的に保証されることになった。その結果、2001年センサスは、行政情報に関して、前二回のセンサスに比べてより広範な原情報の活用が実現した。

一方、作表方法や推計方法の改善としては、まず、90年代後半に社会人口関係の各種の登録情報を個体ベースで相互にリンクし micro-integration によりデータとして統合したデータベース SSD が構築された点がある。その結果、1981年、1991年センサスで単にPRの登録情報と調査データを組み合わせることで作表してきたセンサス結果表が、新たに個体レコードに基づく作表へとその作成方法を質的に転換できた。なお、利用可能な行政情報の範囲が拡大したことは、マッチングや micro-integration 精度そのものの改善にも大きく寄与した。また、2001年センサスでは、RWを新たに導入することで、結果表の数値間の整合性をはかるための一層の努力が払われた。

オランダでは、2004年からの新統計法の施行に伴い、オランダ統計局は行政記録の全面的な統計への使用権限を付与されることになった。また SSD を世帯調査のフレームとして使用することで、調査結果データのより完璧な SSD とのリンクが図られ、さらに行政記録についても ID 番号の整備によりその情報を SSD に組み込む体制がいちだんと整備されることになった。

オランダの virtual census について

とはいえ、virtual census にも問題がないわけではない。なぜなら、virtual census では人口把握の基礎情報を PR に依拠していることから、PR の精度がそのまま virtual census の把握精度を制約することになるからである。

行政情報の場合、事由発生の時点と届出、さらには届出の受理からオランダ統計局への送付の間には一定のタイムラグが発生する。このような届出行為あるいは行政事務遂行に起因する誤差に加えて、PR には、登録情報に宿命的な様々な実態との乖離要因が付き纏っている。例えば、学生の場合しばしば居住地と登録地は異なり、転居者についても届出の遅滞は日常的に発生する。国外への移動者の中にも届出を怠る者がいる一方で、4.6 万人～11.6 万人ともいわれる非登録の不法在留者は、経済活動には参加しているものの PR では登録漏れとなっていることからセンサスの結果数字には反映されていない。その他にも、意図的な不正に基づく現実と登録との乖離もある。例えば、実際には同居しているにも拘らず、社会保障給付受給のために形式的に別の住所で登録しているケースの存在なども確認されている。このような登録における把握漏れあるいは実態と異なる登録行為に起因する人口把握の誤差は、virtual census に固有の誤差というよりはむしろ人口把握の主要情報源を PR に求めるレジスターを基盤とするセンサスに共通するものといえる。

オランダの virtual census が PR をセンサスの基盤情報としているとはいえ、センサス項目の全てを登録情報から入手できるという状況にはない。既存の行政情報から入手できない項目については、それについての行政情報が整備されるまでの間は、現在のような標本調査データに部分的に依拠せざるをえないと考えられる。

むすび

人口センサスについては、その結果数字が政治あるいは行政の根拠数字として重要な社会的意義を持つだけでなく、統計作成においてもいわゆる母集団情報を提供する調査として、他の調査を根拠付ける基盤情報を提供してきた。1960、70 年代以降累進的に深刻化する調査環境の悪化は、統計調査における把握度の低下をもたらした。それは、標本調査だけでなくセンサスについても例外ではない。このため世界の多くの国がセンサスの精度改善のための様々な試みを模索してきており、それは今日、センサス方法の多面的展開となって現れている。

ここで、今日の世界におけるセンサスの多面的展開を試論的に図式化してみよう。センサスの形態的特質を表現する主軸の一方の極には、日本の国勢調査に代表される伝統的な実査に基づく全数調査としての人口センサスがある。一方、北欧諸国などに見られるレジスターベース

の人口把握がその軸の対極をなす。これらのセンサス類型を主軸の両極として、その中間にはレジスターと調査データのいずれをより基盤の情報とするかに応じて多様でしかも副次的展開軸として標本調査の適用形態によって異なる様々なセンサス方法、すなわち、イギリスの one number census やフランスの rolling census, それにドイツのレジスター情報をベースにしたセンサス方式が主軸上のスペクトラムを構成している¹⁶⁾。

PR という登録情報を骨格とするオランダの virtual census は、現在ドイツでその実現に向けての取組みが継続されているセンサスとともに、主軸上では北欧型センサスの亜種、すなわち比較的広範に登録情報が使用可能であるものの全てのセンサス項目を登録情報でカバーする状態にはない広義のレジスターを基調とするセンサス群の一要素を構成しているものと考えられる。しかしながら、北欧諸国、ドイツ、それにオランダのセンサスは、センサス群の中でどのような副次的説明軸を持ち、相互にどのような位置関係にあるのであろうか。レジスター・ベースの統計作成システムを確立したとされる北欧諸国においても、行政情報によって把握できない事項については統計調査が依然として実施されている [(10) p.4]。登録を補完するものとしての統計調査の位置づけなども含め、その点についての検討は今後の課題としたい。

16) Eurostat の Aarno Laihonен は、欧州各国におけるセンサスの実施状況を次の4つのタイプに類型化している。

伝統的センサス実施国	ギリシア、アイルランド、イタリア、ポルトガル、スペイン、イギリス
レジスターベースのセンサス実施国	デンマーク、フィンランド、アイスランド、ノルウェー、スウェーデン
伝統的センサスからレジスターベースのセンサスへ移行中の諸国	オーストリア、ベルギー、ルクセンブルグ、スイス
その他の実施方向（登録と調査の結合、地域あるいは時点ローテーション等）を模索している国	フランス、ドイツ、オランダ

Aarno Laihonен, European Union and Population and Housing Censuses around the Year 2001

オランダの virtual census について

参 考 文 献

- (1) Statistics Netherlands (1999), Informed Consent: Buzzword or Panacea, an invited paper submitted to the Joint ECE/Eurostat Work Session on Statistical Data Confidentiality, Thessaloniki, Greece, 8-10 March 1999.
- (2) Pieter Everaers and Paul van der Laan, The Dutch Virtual Census, <http://unstats.un.org/unsd/dnss/docViewer.aspx?docID=249#start>
- (3) Paul van der Laan (2000), The 2001 Census in the Netherlands-Integration of Registers and Surveys, paper prepared for the conference on 'The Census of Population: 2000 and Beyond' organized by the Cathie Marsh Centre for Census and Surveys Research, 22-23 June 2000
- (4) M.Houbiers (2002), Towards a social statistical database and unified estimates at Statistics Netherlands, *Journal of Official Statistics*, Vol.20. No.1
- (5) Bart F. M. Bakker (2002), Statistics Netherlands' Approach to Social Statistics: The Social Statistical Dataset, *The Statistics Newsletter for the extended OECD Statistical Network*, No.11
- (6) Pieter C. J. Everaers and Paul van der Laan(2003), The Dutch System of Social Statistics: Micro-Integration of Different Sources. ESA/STAT/AC.88/06
- (7) Eric Schulte Nordholt, Marijke Hartgers and Rita Gircour eds. (2004), *The Dutch Virtual Census of 2001-Analysis and Methodology*. Statistics Netherlands, Voorburg
BED6-3607514888AD/0/b572001.pdf
- (8) Eric Schulte Nordholt(2004), The Dutch Virtual Census of 2001-A New Approach by Combining Different Sources, ECE Census meetings, Geneva, 22-26 November.
- (9) The Dutch Virtual Census of 2001: Interview with Eric Schulte Nordholt, ISI Newsletter, Vol. 28, No.3 (84) 2004
- (10) Eric Schulte Nordholt (2007) Record matching for census purposes in the Netherlands, paper submitted for the Conference of European Statisticians, Astana, 4-6 June 2007
- (11) Gosse van der Veen (2007) Changing Statistics Netherlands-Driving Forces for Changing Dutch Statistics, paper presented at the Seminar on the Evolution of National Statistical Systems, Commemorative Event for the 60th Anniversary of the United Nations Statistical Commission
- (12) The Dutch Virtual Census of 2001, IPUMS International, https://international.ipums.org/international-action/sourceDocument.do?document=source_doc_nl2001_tag.xml

[付 記]

本稿は、平成20年度科学研究費補助金(基盤研究C)「センサス機能の変質、新展開およびその統計制度、統計体系への影響に関する総合的研究」(課題番号19530188)の研究成果の一部として公刊するものである。なお、本年8月に実施した海外調査において virtual census の実務に関して貴重な情報提供をいただいたオランダ統計局(Statistics Netherlands)の社会・空間統計部門上席研究員・プロジェクトリーダーの Eric Schulte Nordholt 氏にこの場を借りて謝意を表しておきたい。

Summary

On the Dutch Virtual Census

The history of modern enumeration based population census seems to have entered into the new stage characterized by wide dispersion in its practices. Latest developments in world population censuses, such as register based practices in Nordic countries, sampling assisted British one number census, French trial of rolling census and German register oriented census, demonstrate a multi-dimensional range of spectrum among traditional enumeration, register based censuses and sampling assisted censuses.

This paper aimed to elucidate the characteristics of Dutch experiences called virtual census; a new approach for obtaining a set of census results through micro-based linking and integrating register and survey data. Statistics Netherlands first introduced virtual census in 1981 to substitute traditional enumeration based censuses after the nation wide dissension at the 1971 census. The 2001 virtual census was an improved version of preceding two decennial censuses which had many inconsistencies in their results. The Dutch methodology formulated in the virtual census illustrates a new combination of register and survey records for census purposes under the tough resource constraints. The Dutch virtual census may contribute as a new element to analyze census spectrum and to foresee the future possibilities of census taking.