

# オーストラリアにおける公的統計の二次的利用について

——オンデマンド集計システム TableBuilder を中心に——<sup>†</sup>

伊藤 伸 介\*

谷道 正太郎\*\*

小島 健 一\*\*\*

1. はじめに
2. オーストラリアにおける公的統計マイクロデータの提供
3. オーストラリアにおけるオンデマンド集計システム——TableBuilder——
4. オーストラリアのTableBuilderにおける秘匿処理の考え方について
5. むすびにかえて

## 1. はじめに

公的統計のマイクロデータにおいては、①匿名化マイクロデータ（個票データに匿名化処理が施されたデータ）の提供、②個票データの提供、③オーダーメイド集計、④オンデマンド型の集計サービス（リモート集計）といった様々なチャンネルが存在する（伊藤（2016））。一方、わが国の統計法は、公的統計データの二次的利用の促進を図ることを踏まえ、2007年に改正され、2009年に全面施行された。現行の統計法においては、学術研究等を主要な対象として、調査票情報の提供（統計法33条）、オーダーメイド集計（統計法34条）、匿名データの提供（統計法36条）の二次的利用が可能となっている。

一方、わが国では、利用者が調査項目（表頭・表側）を選択するだけで集計結果を自動的に出力

---

<sup>†</sup> 本稿は、谷道・伊藤・小島（2016）をもとに、オーストラリアにおける公的統計の二次利用の最新状況について加筆・修正をしたものである。共著者である谷道正太郎氏（（独）統計センター経営審議室課長代理）と小島健一氏（（株）野村総合研究所）の了解をいただくだけでなく、（独）統計センターより取りまとめの許可をいただいた上で、伊藤が筆頭著者として本稿を作成した。関係各位にお礼を申し上げたい。また、本稿の旧稿の作成においては、（独）統計センターの「政府統計オンデマンド集計等に関する調査研究」の一環として、本稿の共著者である伊藤と小島、および小林慎太郎氏（（株）野村総合研究所）がオーストラリア統計局の職員に対して行ったヒアリング調査（2016年2月22日～24日）に基づいている。

する、オンデマンド型の集計サービスについて関心も少なくないように思われる。平成26年3月25日に閣議決定された第Ⅱ期「公的統計の整備に関する基本的な計画」では、統計データの有効活用の推進を図る観点から、オンデマンドによる統計作成に関する技術的検証等について実用化に向けた検討を行うことが言及されており、オンデマンド型の集計機能サービスについても、その実用性に関する議論が展開されている。また、平成27年7月に開催された「第20回統計データの二次的利用促進に関する研究会」においては、「統計データの二次的利用の検討状況（全体像・イメージ）」が提案され、オーダーメイド集計に関しては、「利用制限の緩和、『オンデマンド集計』の実用化」が今後の課題として指摘されている。

公的統計においては、集計結果表が公表される場合、調査客体が特定されないような秘匿措置が求められるが、一般には、結果表が詳細になるにつれ、また、関連する結果表が増えるにしたがって、秘匿の度合いや困難さは増大する。ゆえに、オンデマンド集計において利用者の様々な要求に応じて集計表を作成・提供するためには、作成された集計表の秘匿に関する十分な検討が求められる。

他方、諸外国においては、公的統計データの提供形態の1つとして、オンデマンドによる集計システムが展開されている国が存在する。これらの国々では、集計表の有用性と秘匿性に関するトレードオフ関係を何らかの方法で克服し、オンデマンド集計の提供を実現させている。例えば、欧州統計局（Eurostat）で提供されている Census Hub、アメリカの American Fact Finder 等が知られているが、オーストラリアで開発された TableBuilder は、先進的な取組み事例の1つとして国際的にも注目されている。TableBuilder においては、インターネットを介してシステムにアクセスし、調査項目を指定することによって、自動で集計された統計表をリアルタイムに受け取ることが可能になる。

そこで、本稿では、オンデマンド方式による統計作成・提供を実現しているオーストラリアを対象に、オンデマンド集計システムである TableBuilder を中心に、オーストラリアにおける公的統計の二次的利用の現状を明らかにする。また、TableBuilder については、法律上の側面や技術的な側面を明らかにするだけでなく、実現に寄与していると思われる運用面についてもその特徴を考察することにした。

## 2. オーストラリアにおける公的統計マイクロデータの提供

諸外国では、秘匿性と利用者のニーズを踏まえた形で、マイクロデータの提供に関する多様なチャンネルが存在する。表1は、諸外国における公的統計の二次的利用について類型化を行ったものである。表1を見ると、統計データの二次的利用の種類に関しては、個票データ（confidential data（あるいは非識別データ、deidentified data））への直接的なアクセスの可能性および分析結果の提

表1 諸外国の政府統計の二次的利用の種類

二次的利用の種類		特 徴	状 況
直接利用型	オンサイト型	・調査実施者の指定する施設内において、調査実施者が提供するデータを利用することができる。	わが国において、すでに実施済みである。
	直接提供型	・利用者の研究室等、特定の施設以外の場所で、一定の条件のもとに調査実施者が提供するデータを利用することができる。	
	リモート提供型 (リモートアクセス)	・リモートではあるが、直接統計局等のサーバにアクセスし、データ分析等を行い、結果表を受け取ることができる。	わが国では実現していない。
プログラム送付 集計型	参照可能型	・ネットワークを経由して調査実施者が提供するデータを閲覧した上で、作成したプログラムを送付し、集計された結果を受け取ることができる。	
	参照不可能型	・ネットワークを経由して調査実施者が提供する参考情報を利用した上で、作成したプログラムを送信し、集計された結果を受け取ることができる。	
オーダー メード型	従来型	・調査実施者に対して集計の委託を行い、調査実施者によって作成された集計表を受け取ることができる。	わが国において、すでに実施済みである。
	リアルタイム提供型	・インターネットを介してシステムにアクセスし、集計項目を指定することによって、自動的に作成された集計表をリアルタイムに受け取ることができる。	わが国では実現していない。

出所 「諸外国における二次的利用の現状について」（総務省）を一部加筆・修正。

供形態によって、類別化が可能である。表1によれば、公的統計の個票データの二次的利用については、直接利用型、プログラム送付集計型とオーダーメード型に大別される。そして、わが国において現在可能なのは、直接利用型の中で、①利用者の研究室等、特定の施設以外の場所で、一定の条件のもとに調査実施者が提供するデータを利用するタイプである「直接提供型」、②調査実施者の指定する施設内において、調査実施者が提供するデータを利用する「オンサイト型」、および③オーダーメード型の中で、調査実施者に対して集計の委託を行い、調査実施者によって作成された集計表を受け取る「従来型」である。①については、統計法第33条に基づく調査票情報の提供および統計法33条に基づく匿名データの提供がそれに該当する。②は、統計法第33条において現在試行的に行われているオンサイト施設における個票データの提供が該当する。③は、統計法第34条におけるオーダーメード集計を示している。

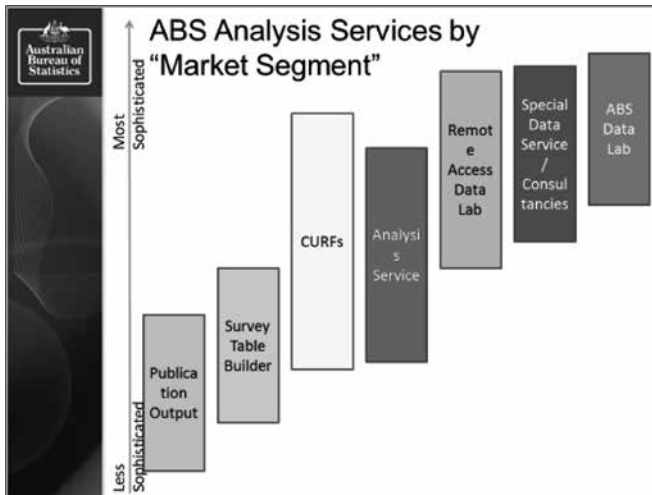
一方、オーストラリアにおける統計の作成および提供のサービスは、図1で概観することができる。オーストラリア統計局（Australian Bureau of Statistics=ABS）は、集計表の作成・公表を行

うだけでなく、匿名化マイクロデータである CURFs (=Confidentialised Unit Record Files) の作成・提供、さらには、プログラム送付による CURFs の利用サービス (Remote Access Data Lab=RADL) やオンサイト施設 (ABS Data Lab) による個票データの利用サービスを行っている。オーストラリアのオンデマンド方式による統計作成については、図1の中で集計結果表 (Publication Output) の隣に位置する TableBuilder が該当する。

CURFs は、学術研究用ファイル (scientific use file) としてのみ提供されており、秘匿の程度に応じて、基本ファイル (Basic CURFs) と詳細ファイル (Expanded CURFs) の複数のファイルが作成されている。CURFs の作成においては、偶発的な個体特定 (spontaneous recognition) の回避という観点から、Elliot 等によって開発された SUDA (Special Uniques Detection Algorithm) (Elliot et al. (2002)) を用いた秘匿処理が行われている。また、オーストラリアの場合、世帯単位で抽出された、抽出率の異なる (1%, 5%) 人口センサスのマイクロデータ (Census Sample Files (CSFs)) が作成・提供されている。なお、CURFs の基本ファイルの提供形態としては、オーストラリアの場合、これまで CD-ROM による提供が行われていたが、2017年10月時点においては、ID とパスワードを取得することによって、ABS の HP から有料でダウンロードすることが可能になっている。

一方、オーストラリア統計局に ABS Data Lab というオンサイト施設が設置されており<sup>1)</sup>、有料による個票データの利用サービスが行われている。ABS 内部に設置されている RADL (Remote

図1 オーストラリア統計局が提供する統計作成サービスの体系



出所 Elazar (2013).

1) Date Lab には、オンサイト施設による個票データの利用サービスだけでなく、(大学の研究室と  
いった場所からの) リモートアクセスによる個票データの利用のサービスが展開されている。

Access Data Lab) では、CURFs の詳細ファイルにリモートでアクセスすることが可能なプログラム送付型の統計作成・提供サービスがなされている。また、ABS Data Lab においては、標本抽出率が5%~10%の人口センサスの個票データにアクセスすることが可能である。なお、表2は、ABSによって行われている各種の公的統計の二次的利用のサービスに関する特徴を示したものである。

オーストラリアでは公的統計データにおける秘密保護は、ABS全体に課されており、それは、「センサスおよび統計法」(Census and Statistics Act 1905、以下『統計法』と呼称)で規定されている。公的統計の秘密保護、統計の公表と提供、センサスに関する情報の開示に関する法律は、それぞれ『統計法』の第7条、第12条と第19A条に明記されている。

## 7 忠誠と機密保護の履行

本法またはその規則に基づいて与えられた権限と義務を遂行する全ての職員は、本法に基づいて職務に入ったり、権限を行使したりする前に、規定の様式に、忠誠と秘密保護の履行を証人の前で署名をしなければならない。

## 12 統計の公表等

- (1) 統計局長 (Statistician) は、本法律に基づいて収集された統計情報を集約 (集計) して分析しなければならない、そしてその集約・分析、又はそれら集計結果や要約 (統計量) を、公表・提供しなくてはならない。
- (2) 細節 (1) に係る集計結果、又は要約は、(1) 特定の個人または組織の識別が可能になるようなやり方で公表・提供してはならない。
- (3) 統計局長は本節の下で公表・提供する集計結果や要約に係る費用を徴収してもよい。

## 19A センサス情報の非開示

### 政府機関へのセンサス情報の非開示

- (1) 現在又は過去に、統計局長又は職員であった者は、当該統計の公表日より99年間いかなる時も、本法に基づく範囲以外で、
  - (a) 当該センサスに関して、第10条に基づき、統計局長又は権限を有する職員に与えられる様式に含まれるいかなる情報も、政府機関へ開示する、又は口外することを要求されてはならない。
  - (b) 自発的にそのような情報を政府機関へ提供してはならない。

### 裁判所や裁決機関へのセンサス情報の非開示

- (2) 現在又は過去に、統計局長又は職員であった者は、当該統計の公表日より99年間いかなる時も、本法に基づく範囲以外で、

表2 オーストラリア統計局が提供する公的統計の二次的利用サービスの比較

提供されるデータの種類の種類	TableBuilder	匿名化マイクロデータ基本ファイル (Basic CURF)	匿名化マイクロデータ詳細ファイル (Expanded CURF)	詳細マイクロデータ (個票データ, Detailed Microdata)
提供形態	Table Builder	Microdata Download	Remote Access Data Laboratory (RADL)	DataLab (オンサイト) DataLab (リモートアクセス)
利用方法	安全な web 上でログイン	安全な web 上でログインし、ファイルをダウンロード	安全な web 上でログイン	オーストラリア統計局のオンライン施設でアクセス
アクセス可能なデータの粒度	集計データ	マイクロデータ	マイクロデータ	マイクロデータ
利用するデータの粒度	非常に詳細なデータ	やや細かいデータ	詳細なデータ	非常に詳細なデータ
環境				
利用者自身の PC 環境からアクセス可能	○	○	○	○
利用者自身の環境から個体レコードの分析が可能	×	○	×	×
将来の利用のために分析結果の保存が可能	○	○	○	○
要求される技術とソフトウェア				
インターネットのアクセス	○	○	○	○
要求される分析用のソフトウェア	×	○	×	×
分析用のプログラムを用いた経験	×	×	○	○
研究のための利用				
中間的な成果物	○	○	○	×
基本統計量 (平均や範囲を含む)	○	○	○	○
集計表	○	○	○	○
グラフによる結果	○	○	×	○
地図	○	×	×	×
最頻値	×	○	○	○
集中度	×	○	○	○
分布に関する高次のモーメント	×	○	○	○
線型回帰モデルの係数	×	○	○	○
非線形回帰モデルの係数	×	○	○	○
推定値における要約統計量と検定統計量	×	○	○	○
利用の条件				
オーストラリア統計局の職員による分析結果の露見に関するチェック	×	×	×	○
利用料金の請求	○	○	○	○
データの利用に関する訓練の必要性	×	×	×	×
システムの運営面における利用者の制限	○	×	○	○
法律の適用による利用者の制限	×	○	○	○
利用条件 (Terms and Conditions of Use) に基づく利用者の制限	○	○	○	○

なる時も、

(a) 当該センサスに関して、第10条に基づき、統計局長又は権限を有する職員に与えられる様式に含まれるいかなる情報も、裁判所や裁決機関へ開示する、又は口外することを要求されてはならない。

(b) 自発的にそのような情報を裁判所や裁決機関へ提供してはならない。

(以下省略)

法第7条は、ABSの職員における守秘義務を明記したものである。法第12条の第1項では、「統計局長は、本法律に基づいて収集された統計情報を集約（集計）して分析しなければならず、そしてその集約・分析、又はそれら集計結果や要約（統計量）を、公表・提供しなくてはならない」ことが明記されている。また、法第12条の第2項では、「集計結果、又は要約は、特定の個人または組織の識別が可能になるようなやり方で公表・提供してはならない」ことが述べられている。これは、ABSで公表されている集計結果表について、攪乱的手法の適用を含む秘匿処理を必要とする法的な根拠となっているだけでなく、オーストラリアにおいて個票データや CURFs の提供を可能にするための規定だと言うことができる（勝浦・小林（2006））。さらに、法第12条は、TableBuilderにおいてオンデマンドで作成された集計表における秘匿処理の必要性を表す条文になっている。したがって、法第12条は、オーストラリアにおいて公的統計の二次的利用に関する基本的な条項でもあると言えよう。

法第19A条は、センサス情報の非開示に関する条文である。オーストラリアでは、第二次世界大戦時に、ヨーロッパからの移民を多く受け入れてきた経緯もあり、公的統計の守秘義務に対して強い規制を強いてきた。とりわけセンサスにおいては、『統計法』で規定された法第7条、第12条と第19A条の遵守が、強く求められている。人口センサスについても、CURFsや個票データという形で提供されていることを勘案すると、このような法的根拠に基づいて、オーストラリアにおける公的統計の二次的利用のサービスが展開されたと考えることができる<sup>2)</sup>。

---

2) センサスに関しては、オーストラリア統計局の役職員に対して厳しい法令遵守が求められている。その内容は、TableBuilder User Manualで紹介されている。その内容の一部は、以下のとおりである。

- ・本法制度の対象者は、常勤と非常勤のいずれにも該当する。
- ・間接的直接的にかかわらず、この法のもとで集められた、どのような秘匿情報についても漏洩した現在の職員または過去の職員に対し、違反と見なす。
- ・違反のペナルティは、最大5,000ドル（マニュアルによる記載）の罰金、二年以内の禁固のどちらか一方、または両方である。
- ・違反者は、オーストラリア統計局から懲戒される。
- ・オーストラリア統計局は Privacy Act 1988に基づく原則を遵守する義務を負う。

### 3. オーストラリアにおけるオンデマンド集計システム——TableBuilder——

オーストラリアのオンデマンド集計システムである TableBuilder は、マイクロデータから統計表、グラフを作成するためのオンラインツールである。表頭と表側の変数を選ぶことによって、統計表やマップを簡単に作成することができるのが、TableBuilder の特徴である。なお、Table Builder については、2009年より、サービスが本格的に開始されている。

TableBuilder で利用可能なサービスは、センサスデータ (Census) とそれ以外のデータ (Others) に関する集計サービスに大別されるが、それは表3のように体系づけられる。基本的な構成は、TableBuilder Pro と呼ばれる有料のサービスと TableBuilder Basic と呼ばれる無料のサービスの2種類であり、アクセス可能な公的統計データもセンサスのデータとそれ以外のデータに類別される。また、公的統計データの利用可能な年次 (あるいはセンサスの実施時期) によって、異なるサービスが行われる。

サービスの体系がセンサスのデータとそれ以外のデータに分けられているのは、ABS の統計データを管轄している部局が異なるからであって、秘匿処理の違いといった技術的要素によるものではない。ただし、ツールの機能から見ると、センサスのデータとそれ以外のデータでは、機能が異なっている。センサス向けの TableBuilder では、地図作成機能を使用することが可能であるが、センサス以外の TableBuilder においては、標本調査のデータセットから算出された基本統計量 (ウェイト, RSE, 連続データ項目, 平均, カスタム範囲, 分位点を含む) のみが、算出され、提供されている。

---

また、センサスに関しては、情報の収集から保存に関するプロセスがマニュアルで詳細に定められることによって、その秘匿性が確保されている。具体的には、以下のとおりである。

(情報の収集に係るプロセスの規定)

- ①センサス収集者は、データフォームを守るために、常時、収集ルートが安全であることを責任を負う。
- ②オーストラリア統計局 (ABS) は、センサスデータが収集者によって見られることを懸念する世帯に対し、1) プライバシー封筒の提供、または、2) eCensus での回答の選択肢を与える。
- ③疑心の強い世帯は、センサス収集者に、データが ABS に到着したことを照明するメールバックを求めることができる。

(収集した情報の処理・保存に係るプロセスの規定)

- ① Collection Center に集められたデータは、Data Processing Center に、安全な (セキュアな) 状態で移される。
- ②常時、警備を行うことによって、権限の無い人の Processing Center への侵入が防止される。
- ③処理された後の書類は、オーストラリア統計局職員の立会のもとで溶解処理される。
- ④処理後、データとしての保管は、ABS computer system で実施される。日常的にパスワードは変更され、アクセスコントロールや認証の形跡もチェックされる。



表3 TableBuilderの種類

	人口センサス	その他のデータ
有料	2006年, 2011年と2016年の人口センサスに関する TableBuilder Pro data series	人口センサス以外のデータに関する TableBuilder data series
無料	2006年, 2011年と2016年の人口センサスに関する TableBuilder Basic	—

注 TableBuilder Basic 2006とは、以前、CDATA Online 2006として提供されていたサービスの後継となる集計システムである。

出所 <http://www.abs.gov.au/websitedbs/censushome.nsf/home/Tablebuilderdetails?opendocument?navpos=240>  
<http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Microdata+prices>

表4 TableBuilder Pro と TableBuilder Basic の違い

	TableBuilder Basic	TableBuilder Pro
対象者	• センサスの利用経験があり、基本的な統計表を作るスキルを持つ者	• Basic に比べてさらに進んだスキルを持つ者
価格	• 無料	• 有料
対象となるデータベース	• 大半の Census Topics に含まれる項目	• ほぼ全ての Census Dictionary に含まれる項目
操作	• 変数の分類区分が予め設定された統計表の利用が可能	• 利用者の DB やシステムのインポートをし、構築したい統計表の抽出が可能
得られる表の例	• センサスの表の構築や加工 • 加工された地域区分とデータの分類区分の編成と保存 • 各種フォーマットにおけるグラフや図の閲覧や抽出等	• 異なる地域間のデータ比較 • 特定のニーズに適合した、加工された地域区分やデータの分類区分の編成と保存 • センサスデータに関する表の抽出と加工、グラフと関連する地図の作成等

Census Topic	Census Dictionary
<ul style="list-style-type: none"> <li>• 文化・言語の多様性</li> <li>• 教育および学歴</li> <li>• 雇用, 収入と無償労働</li> <li>• 障害者, 要支援者</li> <li>• 子供, 保育</li> <li>• 通常の住居, 国内の人口移動</li> <li>• 人民関連</li> <li>• 持ち家の状況</li> <li>• 持ち家以外の状況</li> <li>• 家族構成・特徴</li> <li>• 住居の特徴</li> <li>• 世帯収入および支出</li> </ul>	<ul style="list-style-type: none"> <li>• 通常の住居に関する継続的な情報</li> <li>• 人民に関する継続的な一連の情報</li> <li>• 雇用者および勤務地に関する継続的な情報</li> <li>• 家族および通常の住居に関する継続的な情報</li> <li>• 住居に関連する継続的な一連の情報</li> </ul>

出所 <http://www.abs.gov.au/websitedbs/censushome.nsf/home/Tablebuilderdetails?opendocument?navpos=240>

つぎに、表4は、センサデータを対象としたTableBuilder ProとTableBuilder Basicに関する特徴をまとめたものである。TableBuilder ProとTableBuilder Basicの場合、課金をするか否かという大きな違いがあるが、両者のいずれも、センサスの全てのデータを分析対象とすることができる。ただし、利用できる機能と利用可能な項目に違いがあることから、課金の有無が生じていることがわかる。特に、TableBuilder Basicでは、集計表に用いる変数は、予め設定された分類区分のみが利用可能であって、それらの変数を用いて作成された統計表がオンデマンドで提供可能であるのに対して、TableBuilder Proにおいては、利用者がデータベースの中から希望する変数や分類区分を探索的に選定した上で、統計表を作成することが可能である。

なお、TableBuilder Basicの場合、国外から利用することも可能であるが、TableBuilder Proについては、国内に居住する者に限定されている。また、TableBuilder Proの利用者は通常、自身が所属する組織・団体を介して利用の申請を行うが、ABSは、その利用申請の段階で、組織・団体の確認を行っている。したがって、国内外を問わず、組織の構造等の観点から、ABSがTableBuilder Proの利用組織を判断しているものと推察される。

TableBuilderでは、2017年10月時点で、主として公的統計に関する71個のデータファイルを利用することができる。この71個のデータファイルを見ると、その多くが世帯調査に係る公的統計であること、さらには、行政記録に基づく統計も含まれていることがわかる。

TableBuilderにおいて利用可能な統計の選定は、計画的になされたものではなく、各統計担当部局の裁量によっている。TableBuilderの担当者は、各統計担当部局に、TableBuilderで利用できる統計データの提供を要請しているが、その判断は、各統計担当部局に委ねられている。最終的には、各部局の判断で、提供される統計データファイルが決定される。その意味では、71個のデータファイルは、利用者のニーズに対応したものでなければ、行政側の業務上の利用者に対する還元というものでもないと言える。

TableBuilderによるオンデマンドの集計表の提供については、その大半が、人口・世帯系のデータに基づいている。なお、ABSの担当者によれば、将来的には、度数表（frequency table）ではなく、量的な（Magnitude type）データを含む経済系のデータの提供を考えているとのことで、現在、そのための調査研究がABSで行われている。

TableBuilderの利用者数と利用目的は、表5で示される。2016年2月時点におけるTableBuilder ProとTableBuilder Basicに関する利用者数はそれぞれ、約10,000人と85,000人である。筆者らがヒアリング調査を行った2016年2月時点においても、利用者は増え続けており、Basicの利用者数は1ヶ月に220名、Proは110名程度増加しているとのことである。TableBuilder Proの場合、利用者は、主として研究機関や政府・地方公共団体であり、その利用目的は、地方公共団体の場合、人口や年齢の分布特性の把握であるが、大学・研究機関では、教育目的のためにTableBuilder Proが利用されている。

表5 TableBuilderの利用者数と利用目的

	利用者	利用者のタイプ	利用目的
Census			
	TableBuilder Pro	10,000人 (290機関)	<ul style="list-style-type: none"> <li>研究機関（スタッフ・学生）</li> <li>政府・地方公共団体 ※政府機関は、連邦政府、州政府</li> </ul>
	TableBuilder Basic	85,000人	<ul style="list-style-type: none"> <li>地方公共団体では人口や年齢の概略把握に利用している。</li> <li>研究機関では教育の一環としても利用している。</li> <li>大学が教育目的で利用する際、費用面での優位性が効く。</li> </ul>
	Census 以外	8,000人 ～ 10,000人	<ul style="list-style-type: none"> <li>民間企業を含む様々な機関</li> <li>様々な機関</li> </ul>
			<ul style="list-style-type: none"> <li>様々な目的で利用されている（民間企業ではマーケティング等に用いるケースもある）。</li> <li>Census データ以外においては、地域区分が粗すぎるため、地方政府にはあまり利用されない。</li> </ul>

注 本表は、2016年2月にオーストラリア統計局（ABS）に対して行ったヒアリング調査をもとに作成した。なお、聴取対象によって、利用者数に対する認識が異なっていたが、表5では、一例としての回答を記載している。

TableBuilderの利用者が1万人規模で存在する理由として、ABSは、TableBuilderが表、グラフや地図を作成するシンプルなツールであるだけでなく、センサスについては、TableBuilderにおいて集計の対象となっているレコード数が、他の二次的利用に係るサービスと比較して大きいことを指摘している。先述した通り、ABSは、二次的利用に係る様々なチャンネルを提供しているが、それによって提供されている公的統計マイクロデータは、標本データであることが少なくない。その一方で、TableBuilderにおいては、人口センサスの全数データを対象にした集計を行うことができるのが大きな特徴である。

TableBuilderを利用するには、利用者がABSに登録をすることが求められる。これは、TableBuilder ProとTableBuilder Basicのいずれにおいても必要である。利用登録の手続きは、以下のとおりである。最初に、利用者は、ABSのウェブサイトにあるRegistration Centerに、自身で登録を行う。登録の際は、氏名、所属組織、電話番号、住所、メールアドレス、パスワード等の入力が必要である。その上で、ABSより、ユーザーIDとパスワードを獲得する。一度、自身のアカウントを作れば、利用者は、全ての無償サービスにアクセスすることができる。なお、利用者がTableBuilder Proを利用したい場合、利用者が所属する機関を介しての登録が、別途求められる。

つぎに、TableBuilderの利用者は、表6に示されるような利用規約を遵守することが求められる。具体的には、IDおよびパスワードに関しては、利用者に厳重な管理を求めており、違反者には罰則も設けている。具体的には、TableBuilderの利用者については、本人のID以外からの利用は認められておらず、また、自身のIDを他人に貸し出すことも禁じられている。

表6 TableBuilder 利用者が遵守すべき規約

名称	内容
Terms & conditions (利用規約)	① 自分の ID, パスワードを他人, または他の組織に提供しないこと. ② サービス利用において保存した図表等は, 自身の責任で管理し, オーストラリア統計局は一切関知しないこと. ③ 組織を介してアカウントを保有する場合, 退職等をした際に, それを申し出ること.
ABS Conditions of Sale	特に著作権について以下の対応をすること. ・ TableBuilder の再生産, 再販, 流通, 機能の一部取り出しや他機能との重ね合わせの禁止. ・ 著作権法 (Copyright Act 1968) の第40条~第43条に該当するレベルでの引用, 過度な抜書きの禁止. ・ 出所における ABS の明記. ・ ABS の定義した用語の書き換えの禁止. ・ データ分析後のデータに対する「ABSによる」との記載の禁止 (自身で加工したことへの明記).

注 ABS Conditions of Sale は, TableBuilder 以外の政府統計の二次的利用にも適用される.

出所 <http://www.abs.gov.au/websitedbs/D3310114.nsf/home/ABS+Conditions+of+Sale>

本人の ID を利用しているか否かの確認は, システムで自動的に行うのではなく, ABS の職員が, 利用者に直接電話して行っている. なお, 電話で確認した結果, 利用者以外の人物が TableBuilder を利用していることが判明した場合には, その者に罰則を適用することとしている.

なお, ABS によれば, 利用者本人及び利用者の所属する機関が利用規約を違反することによって利用が停止されるケースもある. ABS の担当者によれば, 違反者 (初回) に対しては, 3ヶ月の TableBuilder の利用停止, 2回目の違反者には12ヶ月の利用停止措置を行っているとのことである.

TableBuilder のアカウント有効期限は, 最大1年である. 毎年6月30日に, TableBuilder 内のデータが更新されることから, その時点で有効期限は切れることになっている. したがって, 継続の意向を持つ利用者は, 再申請を行う必要がある. なお, 利用料金については, TableBuilder Basic の場合, 先述のとおり料金は無料であるが, TableBuilder Pro については, センサスデータについては \$2,350, それ以外のデータに関しては \$1,180となっている (2017年10月時点).

ところで, TableBuilder では攪乱を用いた表を利用者に提供している. この攪乱的措置の結果, どの程度データに信頼をおけるのかに関して, 周知することが求められる. この点について, ABS はつぎのような対応を行っている.

- ① 利用者にはまず, TableBuilder の利用者ガイド (オンラインで提供され, TableBuilder における利用上の制約を記載している) を読むことを要請している.
- ② 公的統計に関する情報, すなわち, どのようにデータが収集されたか, また, サンプルやデー

タアイテムのリスト等の情報は、個々の TableBuilder のデータセットごとに提供される。

- ③ TableBuilder のウェブサイトでは、「小さい値や RSE（標本標準誤差）の大きい結果に対しては、そのデータに信頼を置かない（“No reliance should be placed on small number cells or cells with large percentage Relative Standard Error (RSE) values”）」という記載によって、利用者に注意喚起をしている

ただし、具体的に「どの数字以下を小さい」とするか、「RSE が大きいとはどの程度か」等を示したものは、TableBuilder において提供されない。オンデマンドで作成された集計表によっては、RSE に関する記載がなされていない場合もあることから、利用者は利用目的と照らし合わせながら、自身の判断で、データに対する信頼の範囲を決めることが求められている。

#### 4. オーストラリアの TableBuilder における秘匿処理の考え方について

『統計法』では、いかなる秘匿情報も公開をさせないように処理を行った上で、結果を公表することが求められている。そのため、TableBuilder は、個票データを集計した「元の集計表」に対して、意図的に「ノイズ（攪乱値）を付加する」ことで、データの秘匿性を担保している。そのため、TableBuilder における秘匿処理の重要な特徴は、TableBuilder の集計システムによって得られる全ての集計表に対して攪乱的措置が施されていることにある。

ABS は、秘匿性の担保の方法として、「わずか」なノイズ（攪乱値）を入れる手法が、利用者にとって、もっとも満足のいく技術であると主張している。この「わずか」な攪乱を行うことによって、真の値から乖離することから、ABS は、TableBuilder の利用に際して、「小さい値や RSE（標本標準誤差）の大きい結果に対しては、そのデータに信頼を置かない」ことについて、利用者に注意を促している。また、個票データから「元の集計表」を作った後でセルにノイズを付与していることから、作成された集計表内のセルの値を積み上げると（「加算性」）、合計値と合わなくなる場合もあることを注意喚起している。

以下では、この攪乱をどのように行っているのかについて述べることにしたい。ノイズを入れることによって、分析の対象となる統計表の中の数値が大幅に変わった場合、利用者にとっての情報損失は小さくないと思われる。特に、表の平均値がずれたり、表のデータの分散・共分散構造が崩れたり、また、「0」という結果が「0でない」結果となってしまったりすることは、分析対象となる集計表の価値を著しく減じることが考えられる。したがって、ABS は、以下を前提とした攪乱的手法を適用している。

- ① ノイズ付与に用いる数値の平均値は 0 である。
- ② 真の結果が 0 である場合、その値にはノイズを入れない。
- ③ 標本調査についても、ウェイトやウェイトをかける前の値にノイズを入れるのではなく、あく

まで集計表に対してノイズを入れる（全数調査データと標本調査データで、ノイズ（攪乱値）の入れ方を極力変えない）。

- ④ 同じ変数群に関する集計表を生成する場合は、常に同じノイズ（攪乱値）を加える。

特に、第4の特徴は、Differencing Attack と呼ばれる秘匿情報へのアクセスに対抗する手段である。攪乱する値が、集計表の生成ごとに異なる場合、何度か集計表を生成し、中身を比較していくことで、集計表の中の小さい数値を判別し、その内容を推測することが可能となる場合もある。この行為に対抗するため、TableBuilder では、同一の集計表を作成した場合には、必ず同じ値のノイズを入れる工夫がなされている。

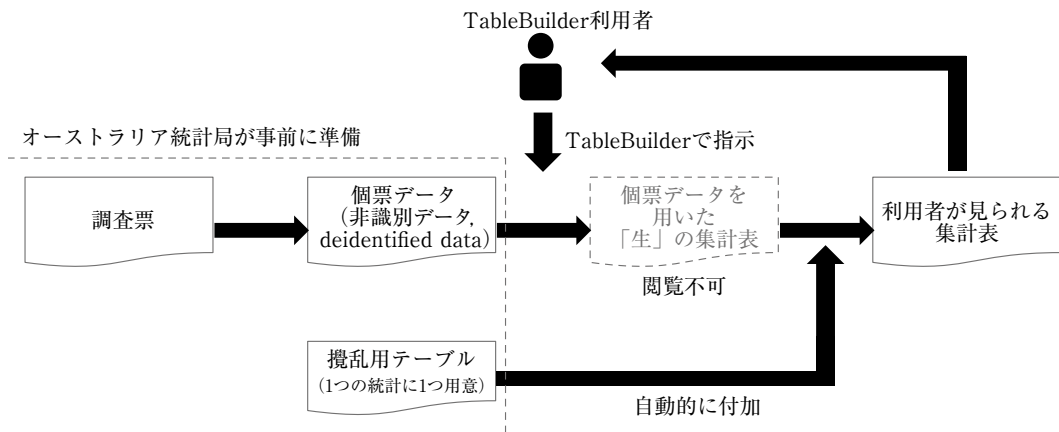
図2は、TableBuilder におけるノイズ付与のプロセスとノイズ付与がなされた集計表の提供のプロセスをまとめたものである。以下では、それぞれの特徴を述べる。

最初に、TableBuilder に用いられるデータであるが、このデータについては、TableBuilder Pro, TableBuilder Basic のいずれも、原データから名前や住所等の直接的な識別子を削除した非識別化（deidentified）された個票データを用いている点は、1つの特徴だと言える。

個票データ自体には、ノイズやスワッピング等の処理はされておらず、また、データキューブのような中間集計表も用いられていない。したがって、幾つかのデータセットは非常に詳細なデータ項目を有する場合もある。ただし、TableBuilder では、最終的な統計表を出力する前に攪乱を入れていることから、安全性に問題は無いと考えられている。

つぎに、攪乱に用いる pTable（Perturbation Look-up Table）と呼ばれる表が、事前に用意されており、1つの統計に1つの pTable が設定される。pTable は、255行×最大30列の行列から構成される。TableBuilder によって作成される集計表が利用者に提供される前に、pTable が適用される。

図2 TableBuilder におけるノイズ付与のプロセス



注 本図は、2016年2月にABS に対して行ったヒアリング調査の内容をもとに作成した。

Andersson et al. (2015) によれば, pTable の作成は以下のように説明される. 最初に, 許容される一連の攪乱 (perturbation) の分布を (1) 式で表すとする.

$$\prod_i \{\pi_{L_i}, \dots, \pi_{U_i}\} \quad (1)$$

そのとき, その攪乱の分布は, 以下の (3) 式を制約条件として, エントロピーの最大になるように決定される ((2) 式).

$$P(B_i = \pi) = - \sum \pi \epsilon_{\pi_i} P(B_i = \pi) \log(P(B_i = \pi)) \quad (2)$$

(制約条件)

$$\begin{aligned} P(B_i = \pi) &\geq 0, \sum \pi \epsilon_{\pi_i} P(B_i = \pi) = 1 \\ i + \pi &\in \{0, l, l+1, \dots\}, l \geq 0 \\ E(B_i = \pi) &= 0 \\ V(B_i = \pi) &\leq v_i \end{aligned} \quad (3)$$

なお, (3) 式における  $l$  は, 統計表に含まれるセルの閾値であり,  $V_i$  は, 許容される攪乱の分散である.

つぎに pTable は, つぎのように設定される. ある統計表をベクトル  $\mathbf{t} = (t_1, \dots, t_k)$  ( $k$  はセルの数) で表す. 各セルに対して, 攪乱の行列  $\mathbf{p} = (p_1, \dots, p_k)$  が pTable に保管されている. このとき, pTable の各列は, セルの度数の一覧表として示されている.  $k$  個のおおのこのセルについて,  $p_k$  は, pTable における  $t_k$  番目の列から取り出される. pTable において対応する行を見つけるために, すべてのレコードは, 固定的な乱数であるレコードキーが割り当てられている. ある統計表において特定のセルの度数が算出されると, その度数に対応するレコードキー (record key) が足し上げられることによって, セルのキー (cell key) が求められる. このセルのキーは, 一般には 32ビットの 2 進数の数字として記述される. これらの数字に基づいて, pTable における行が定まり, それによって, 特定のセルに対する攪乱が取り出される. 統計表  $\mathbf{t}$  に対する攪乱  $\mathbf{p}$  が生成されると, 秘匿処理が施された統計表は,

$$\mathbf{c} = \mathbf{t} + \mathbf{p} \quad (4)$$

で表現される. この攪乱済み統計表の作成においては, 度数表におけるセル度数の分布とは独立に, 特定のセルに対しては同一の攪乱値が付与される.

図 3 は, pTable および集計表に対する攪乱のイメージを示したものである. 攪乱前の集計表において, A1 行 B1 列の度数が 6 であるとする. そのとき, pTable においては, 6 列目の度数の一覧が選択される. つぎに, レコードキーからセルのキーが求められる. 例えば, それが 5 になっ

図3 pTable および集計表に対する攪乱のイメージ

① 攪乱の分布の設定

Perturbation Look-up Table (pTable)	1	2	3	4	5	6	7	...
	1						0	
	2						-1	
	3						3	
	4						-2	
	5						-2	
	6						0	
	.						2	
	.						.	
	.						.	

$$\Pi_6 = \{\pi_{L_6}, \dots, \pi_{U_6}\}$$

(例えば,  $\{-3, \dots, 3\}$ )

・ 分布  $P(B_6 = \pi) \pi \in \Pi_6$  ( $B_6$ : 確率変数) を,

$$\text{エントロピー} = - \sum_{\pi \in \Pi_6} P(B_6 = \pi) \log(P(B_6 = \pi))$$

が最大となるように決定

(制約条件)

$$P(B_6 = \pi) \geq 0 \quad \sum_{\pi \in \Pi_6} P(B_6 = \pi) = 1$$

$$6 + \pi \in \{0, l, l+1, \dots\}, l \geq 0$$

$$E(B_6 = \pi) = 0$$

$$V(B_6 = \pi) \leq v_6$$

② 攪乱値の付与

(Perturbation 前の集計表)

	B1	B2	B3			B1	B2	B3
A1	6	7	5	→ -2だけ攪乱	A1	4	7	5
A2	3	0	2		A2	3	0	2

他のセルに対しても  
同様に攪乱する

Perturbation Look-up Table (pTable)	1	2	3	4	5	6	7	...
	1						0	
	2						-1	
	3						3	
	4						-2	
	5						-2	
	6						0	
	.						2	
	.						.	
	.						.	

	調査項目 A	調査項目 B	調査項目 C	...	Record Key
被調査者①	1	1	1	...	294723402
被調査者②	1	1	3	...	502192840
被調査者③	1	1	2	...	320485
被調査者④	1	1	1	...	66663402
被調査者⑤	1	1	2	...	852284648
被調査者⑥	1	1	1	...	156792411

**Cell Key**  
Record Key を  
“足し算” (mod)  
... → 5  
例えば



たとすると、6列目の度数の列における5行目の数値である-2が選ばれる。この-2がA1行B1列の度数6に付与されて、攪乱済みの集計表におけるA1行B1列の度数は4となる。このような処理が他のセルでも繰り返される。

その一方で、統計実務の観点から見た場合、このpTableの作成には、熟達した技能と経験が必要となっており、秘匿性を担保しつつも、公的統計の収集からなるべく早期にTableBuilderでのサービス提供ができるよう、過去の経験を活用した表となっている。実際、秘匿性の十分な確保を考慮しないのであれば、ABSは、半日程度で自動的に（プログラム等によって）pTableを作成することができる。その一方で、様々な角度からpTableが秘匿性に問題が無いかを検証し、もし問題があれば、pTableを再度作成するという作業を人手で行っていることから、実際には数週間を経て、作成された集計表が利用者に提供される場合もある。

このようなpTableの作成にあたっては、統計実務上の観点から、以下の5点を指摘することができる。

- ① pTableの波形は、0を基軸に左右対称なものである（ただし正規分布とは異なると思われる）。
- ② 左右対称であるため、pTableには0も含まれること（攪乱と言っても、必ず何らかのノイズを入れるのではなく、0が付与される（全く攪乱がされない）ケースもありうる）。
- ③ ノイズの最大値、最小値は、経験則に基づいて決定している。
- ④ 人力で修正する作業を、50~60回程度の試行錯誤によって達成している。
- ⑤ 経験則に則ってpTableを作成しているため、この攪乱的措置が、どの程度、もとの公的統計の情報に損失を与えるのか、どの程度、秘匿に係るリスクが内在したままであるのかに関する評価は実施されていない。

上記の内容は、2006年に、Fraser and Wooten（2006）で展開されたTableBuilderに関する議論とは、大きく異なるものとなっている。Fraser and Wooten（2006）においては、pTableは、攪乱の効果を最大限にするために、数理的かつコンピューターを用いた数値計算によって決定されていた。ABSの担当者によれば、これは、基礎研究と実務との違いであって、論文に記載した内容は、基礎研究として、あくまで学識者等との「コミュニケーション」のためのツールであると位置づけられている。しかしながら、統計実務の面から見ると、pTableの作成方法に関するこのような基礎研究の実務への適用可能性を勘案した上で、pTableを作成していることが考えられる。

最後に、オンデマンドで作成された表に、pTableに含まれるノイズを付与するために、参照用のRecord Keyを使っている点が、攪乱的措置の特徴として指摘される。先述のように、pTableは、255行×最大30列の行列で構成されるが、その中からどの数字を、どのように参照するかに関しては、Record Keyが用いられている。Record Keyとは、個票データ（非識別データ）の1レコードごとに割り振られた値であり、 $2^{32}$ の数字の中からランダムに付与された値である。この値は、1つのマイクロデータに対して一意に決められることから、途中で変わることは無い。利用者が

TableBuilder に対して集計表作成のための指示を行うと、TableBuilder 上で、集計表の中のセルの数値を見ながら、Record Key との対応を図り、この Record Key に対して数学的処理を行うことによって、pTable のどのセルを参照すればよいのかが決定される<sup>3)</sup>。これによって、同じ統計表を生成する場合は、常に同一のノイズ（攪乱値、pTable のセル値）を加えることが担保される。

このような手法を用いて作られた、ノイズが付与された集計表は、安全性が高い集計表とみなされ、ABS の職員等の人手のチェックを介することなく、利用者に提供されている。利用者は、表頭、表側を自由に選ぶことができる。それによって、TableBuilder においては、任意の粒度で集計表を作成することも可能である。多次元のクロス集計表を、探索的に実施できる点は、TableBuilder の重要な特徴と言えるだろう。ただし、運用面から見ると、6～7次元のクロス集計表は、0 が非常に多く入ることから、そうした集計表の提供は困難なものと言える。

一方、地図へのマッピングについても、粒度の細かいレベルでの結果数値の提供が可能だけでなく、安全性も ABS によって担保されている（もっとも、ノイズ（攪乱値）の付与だけでは秘匿が担保できない、または、セルに0しか入らない場合には、統計作成・提供のサービスとして成立しないため、利用者に結果数値の提供はなされない（表頭・表側における変数を選択する際に、変数が選択肢に示されていない）ようにしている）。なお、地図の粒度は、オーストラリア統計地理標準（Australian Statistical Geography Standard = ASGS）に準拠しており、最小の統計的地域区分（Statistical Area Level 1 = SA1）であれば、オーストラリアの54,805地区すべてに関する詳細な地域区分について地図上での分析が可能である。

つぎに、秘匿処理に関する技術的な論点として集計表内の数字の整合性、特に加算性に関しては、以下のように述べることができる（表7）。

先述の通り、TableBuilder は、各セルに対してノイズを入れる手法を採用しているため、基本的にはセルの合計と、予め提示された合計値とは一致しない。センサスデータを含む TableBuilder で使用可能なほとんどのデータセットは加算性の技術を用いていない（2017年10月時点）。しかしながら、加算性の技術が適用されることによって、セルの行と列に関する総和が保たれるように調整が実施され、この加算性に係る整合性を担保するような処理がなされているデータセットも存在する。その場合でも、加算性は特定の1つの集計表の中でのみ担保されることから、例えば上述した SA1 のデータは、SA2（SA1 よりも粗い地域区分）のデータと整合しない。

なお、集計表に対して攪乱を行っている時点で、数字の加算に係る整合性を厳密に図る必要はないのではないかという疑問も生じる。この点については、オーストラリアの文化的な環境とも関係があるものと推察される。具体的には、最低限、総計のみでも整合していればよいと考えられてい

---

3) この数学的処理には、高度な数学的な方法が用いられていることから、詳細については、Thompson et al. (2013) 等を参照されたい。

表7 加算性に関する整理

加算の整合性	あり	なし
秘匿処理の仕方	<ul style="list-style-type: none"> <li>• テーブルの各セルに、ノイズ(攪乱値)を挿入する.</li> <li>• セルの行列の総和が保たれるように、さらに調整を実施する.</li> </ul>	<ul style="list-style-type: none"> <li>• テーブルの各セルに、攪乱を挿入する.</li> </ul>
留意点	<ul style="list-style-type: none"> <li>• 1つの表の中で、データの整合は取れている.</li> <li>• ただし、整合はあくまで1つの表の中に閉じており、調整は、それぞれの表でなされるため、異なる地理レベル等のデータ間での整合は取れていない.</li> </ul>	<ul style="list-style-type: none"> <li>• 表の中で、データの整合が取れていない.</li> </ul>

出所 <http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/by%20Subject/1406.0.55.005~User%20Guide~Main%20Features~Confidentiality~100#ADDITIVITY>

る可能性もある。

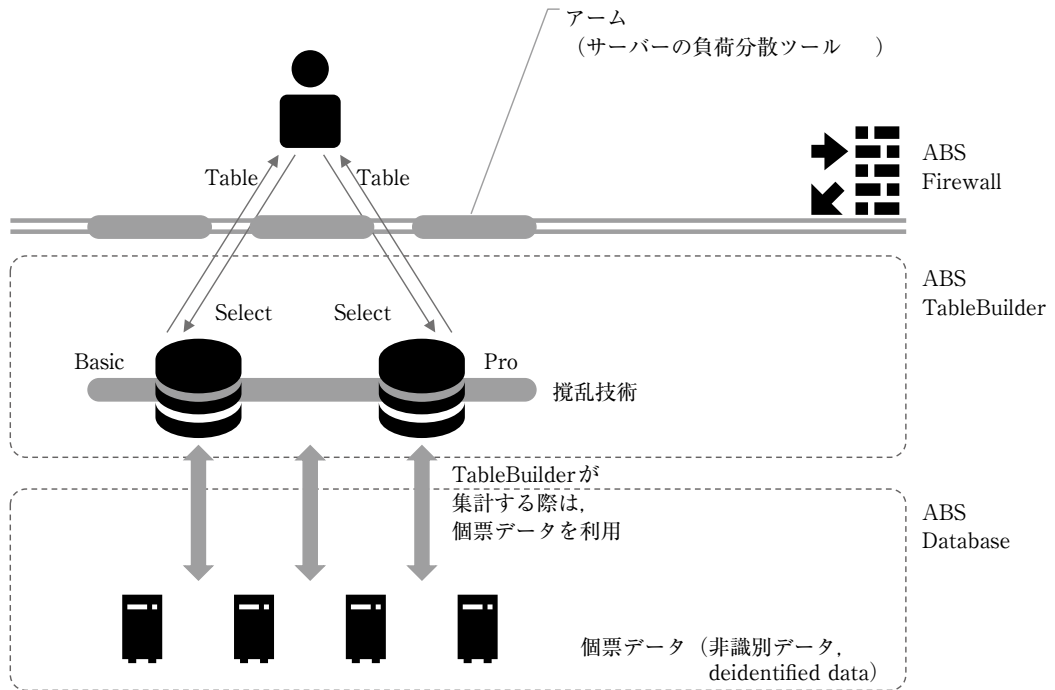
このように、TableBuilderの秘匿処理技術の特徴を明らかにしてきた。一方で、こうした攪乱的処理を用いた統計データの「正確性」については、わが国とは基本的な考え方が異なることも考えられる。TableBuilderから得られる集計表に対する攪乱的措置の考え方については、つぎの3点にまとめることができる。

- ① 利用者の一部は、攪乱の影響を懸念している。ただし、この処理によって、ランダムな誤差が含まれるものの、集計表における有用性を大きく損ねるものではない。
- ② TableBuilderは、全てのセルにノイズを入れている。ノイズを付加されたデータではあるが、ノイズはあまり大きくない。したがって、統計量としての分布特性は、ノイズを入れたとしても変わらない。
- ③ データセットに対して適用される攪乱の値は一定であるため、得られた結果数値をもって、「真の値」として報告している。

オーストラリアの公的統計においては、法12条第2項に基づいて公表された統計表も含めて、何らかの形で秘匿処理がなされている。このことは、TableBuilderを公的統計の二次的利用における統計作成・提供サービスとして成立させるための重要な要件でもあると言える。オーストラリアでは世帯・人口系の公的統計に対しては、TableBuilderと同様のノイズ(攪乱値)を入れているだけでなく、事業所・企業系の公的統計に対しては、個人情報秘匿性を確保するために閾値を用いた秘匿処理を行う等、別の手法も導入している。また、こうした秘匿処理の中には、偶発的な個体特定を回避するための処理を行うことも含まれる。

つぎに、オンデマンド集計システムの安全性を確保するためのTableBuilderの情報システムの

図4 TableBuilder の情報システムの構成



注 2016年2月にABSで行ったヒアリング調査に基づき作成した。

構成について述べることにしたい。TableBuilder の情報システムの構成は、図4に示されている。その特徴としては、第1に、TableBuilderにおいて集計を行う上で用いているデータは、個票データであることが指摘される。ただし、前述の攪乱的な方法を用いることによって、TableBuilderにおいてはレコードの属性情報は露見されないことから、個人情報（人、世帯、車、事業所、企業等）の秘匿性は全て確保される。それゆえ、個人等のプライバシーが漏洩することがなく、Database層で、TableBuilderに係る特別なセキュリティ対策は実施されていない。同様に、秘匿性が破られることに対する危機管理マニュアル等も作成されていない。

なお、システム構成に関する特徴としては、データベースの構成と、TableBuilderのシステム開発において、Space Time Research社（Space Time Research Pty Ltd.）という民間企業のリソースを活用していることが指摘される。このSpace Time Research社は、ABSと1986年以降密接な関係を構築してきた。TableBuilderの攪乱に関するアルゴリズム自体は、ABSによって開発されているが、システムの開発、コーディングおよび、データベースの設計・開発は、Space Time Research社によって行われている。特にデータベースは、オラクルやSQLサーバのような汎用のデータベース製品ではなく、Space Time Research社によって独自に開発されている。その詳細については、ABSも把握していない。Space Time Research社の開発したシステムによって、公的統計データの高速処理が可能となっており、データ処理の多くは、即時的な対応が可能となっ

ている。ただ、マッピング等で地域区分を細かくした場合には、その限りではない。例えば地図上に500万のセルを表現しようとした場合には、システム負荷が高く、処理に24時間以上要すると考えられている。そのため、TableBuilderの運用にあたっては、500万セルがその上限となっており、それ以上に負荷のかかる処理については、ABSは原則として受け付けていない<sup>4)</sup>。

## 5. むすびにかえて

本稿では、オーストラリア統計局が開発したオンデマンド集計システムであるTableBuilderを中心に、オーストラリアにおける公的統計の二次的利用の基本的な特徴を明らかにした。本稿においては、TableBuilderに関して、その法制度面や運用面での特徴、さらには集計表を出力するための秘匿処理の方法を明らかにした。オンデマンド集計では、利用者のリクエストに応じて集計表を作成・提供することから、集計表の秘匿性をどのようにして確保するかが重要な点だと言える。TableBuilderの場合、集計表における秘匿処理の方法として、出力された集計結果表のすべてのセルにノイズが導入されていることが秘匿性を確保するための基本的な特徴となっている。

なお、TableBuilderに関しては、以下のような方向性がABSによって模索されている。

第1は、TableBuilderの統合である。具体的には、センサスとそれ以外の標本データでは、それぞれ別のタイプのTableBuilderが開発されてきたが、将来的には複数のタイプのTableBuilderの統合が計画されている。

第2は、利用可能な公的統計の範囲の拡大である。マーケティング等、ビジネスに関するデータは、これまでTableBuilderで提供されたことがなかったため、現在、個人や世帯とは異なる秘匿処理の手法を用いて提供することが検討されている。

第3は、複数の統計調査のリンケージによるサービスの提供である。例えば、移民に関するデータとセンサスデータとのリンケージデータもTableBuilderで分析可能になることが計画されている。

こうした、ABSにおけるTableBuilderの動向については、わが国におけるオンデマンド型の提供サービスを議論する上で参考になる点が少なくないように思われる。一方、わが国において、TableBuilderのようなシステムに基づいてオンデマンド集計システムを展開しようとするのであれば、オンデマンド集計システムによって作成される集計表の有用性、とりわけ集計表に含まれるノイズの付与に関する方法論上の整理、さらにはオンデマンド集計システムに保管されるデータの

---

4) システムへの過大な負荷が利用者側のTableBuilderの操作に影響が及ばないように、システム構成には工夫がなされている。具体的には、利用者の操作によってサーバを独占してしまうことが無いよう、図3に示したようなアームを3つ設けることによって、システムにかかる負荷の分散を図っている。

構造（個票データかあるいは高次元の集計表か）等についての議論が必要なように思われる。オンデマンド集計の可能性を検討する上では、このような諸外国における公的統計の二次的利用の動向を踏まえつつ、オンデマンド集計に関するわが国の法制度的な位置づけと秘匿処理の可能性の両面から、具体的な検討を行う必要があるだろう。

#### 参考文献

- 伊藤伸介（2016）「わが国における公的統計のデータシェアリングの現状と課題」『情報管理』, Vol. 58, No. 11, 836-843頁
- 勝浦正樹・小林健太郎（2006）「オーストラリアにおけるデータアーカイブとマイクロデータの提供」『名城論叢』第7巻第2号, 69-82頁
- 谷道正太郎・伊藤伸介・小島健一（2016）「オーストラリアのオンデマンド集計に関する調査研究」『製表技術参考資料』No. 33, 1-33頁
- Andersson, K., I. Jansson, K. Kraft (2015) "Protection of frequency tables – current work at Statistics Sweden", Paper Presented at UNECE/Eurostat Work Session on Statistical Data Confidentiality, Helsinki, Finland, pp. 1-20.
- Elazar, D. (2013) "Innovative Microdata Access-Confidentialising on the fly", UNECE Work Session on Statistical Data Confidentiality, Ottawa, Canada, [http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2013/Topic\\_7\\_ABS.pptx](http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.46/2013/Topic_7_ABS.pptx)
- Elliot, M. J., A. M. Manning, R. W. Ford (2002) "A Computational Algorithm for Handling The Special Uniques Problem", *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, Vol. 10, No. 5, pp. 493-509.
- Fraser, B., J. Wooten (2006) "A Proposed Method for Confidentialising Tabular Output to Protect against Differncing", Paper Presented at UNECE/Eurostat work session on statistical data confidentiality, pp. 1-6.
- Thompson, G., S. Broadfoot, D. Elazar (2013) "Methodology for the Automatic Confidentialisation of Statistical Outputs from Remote Servers at the Australian Bureau of Statistics", Paper Presented at UNECE Work Session on Statistical Data Confidentiality, Ottawa, Canada, pp. 1-37.

[オーストラリア統計局の HP については2017年10月31日現在]

<http://abs.gov.au/websitedbs/censushome.nsf/home/tablebuilder>

<http://abs.gov.au/websitedbs/censushome.nsf/home/tablebuilderdetails?opendocument&navpos=240>

<http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Microdata+prices>

<http://www.abs.gov.au/websitedbs/D3310114.nsf/home/Expected+and+available+Microdata>

<http://www.abs.gov.au/websitedbs/D3310114.nsf/89a5f3d8684682b6ca256de4002c809b/c00ee824af1f033bca257208007c3bd5!OpenDocument>

<http://www.abs.gov.au/websitedbs/d3310114.nsf/home/ABS+Conditions+of+Sale>

<http://www.abs.gov.au/websitedbs/censushome.nsf/home/factsheetsas>

(\*中央大学経済学部教授 博士(経済学))

(\*\* (独)統計センター経営審議室課長代理)

(\*\*\* (株)野村総合研究所 ICT・メディア産業コンサルティング部主任コンサルタント)