

統計の作成、提供等における  
情報通信技術の活用に関する調査研究

報告書

平成 26 年 3 月

株式会社 野村総合研究所



# 目次

<b>第1章 調査研究と結果の概要</b>	<b>1</b>
1 調査研究の目的	1
2 調査研究の方法	2
1) 調査研究の全体フロー	
2) 有識者による研究会の設置	
3) 国内調査・海外調査の視点	
3 調査結果の概要	6
1) 調査実施ステージ	
2) 結果活用ステージ	
<b>第2章 統計調査において活用される情報通信技術の概要</b>	<b>13</b>
1 調査実施ステージ	13
1) データ収集方法の分類	
2) コンピューター支援によるデータ回収(CAI: Computer-Assisted Interview)	
3) 調査実施ステージにおいて活用される情報通信技術	
2 結果活用ステージ	18
1) データ作成・集計機能	
2) 分析機能	
<b>第3章 調査実施ステージにおける情報通信技術の活用</b>	<b>21</b>
I 国内における情報通信技術の活用に関する取組	22
・ 統計局および統計センター	22
・ 国勢調査における取組	
(1) 調査概要	
(2) 情報通信技術の位置付け	
(3) 組織体制・機能	
(4) 情報通信技術の導入経緯・目的	
(5) 情報通信技術の活用に関する評価	
(6) 課題と今後の方向性	
II 海外における情報通信技術の活用に関する取組	29
1 アメリカセンサス局(United States Census Bureau)	29
1) Centurion(インターネットによるデータ収集システム)を活用した取組	
(1) Centurion の概要	
(2) 開発の経緯・体制	
(3) 基本的な機能	
(4) 主な特徴	
(5) セキュリティに関する取組	
(6) 今後の課題	
～Centurion コントロールセンター～	

2	アメリカ労働統計局(United States Bureau of Labor Statistics)	36
	・ CES(Current Employment Survey:雇用状況調査)における取組	
	(1) 調査概要	
	(2) 情報通信技術の位置付け	
	(3) 組織体制・機能	
	(4) 情報通信技術の導入経緯・目的	
	(5) 情報通信技術の活用に関する評価	
	(参考) インターネット回答のインフラストラクチャーIDCF(Internet Data Collection Facility)について	
	(6) 課題と今後の方向性	
3	カナダ統計局(Statistics Canada)	50
	・ LFS(Labour Force Survey:労働力調査)における取組	
<b>第4章 結果活用ステージにおける情報通信技術の活用</b>		<b>55</b>
I	国内における情報通信技術の活用に関する取組	56
1	統計センター(e-Stat)	56
	1) e-Stat の概要	
	2) 取組の概要	
II	海外における情報通信技術の活用に関する取組	61
1	カナダ統計局(Statistics Canada)	61
	1) 取組の概要	
	2) 情報通信技術の導入経緯	
2	シカゴ大学 NORC	62
	1) 取組の概要	
	2) 情報通信技術の導入経緯	
3	ミシガン大学 ICPSR	63
	1) 取組の概要	
	2) 情報通信技術の導入経緯	
<b>第5章 結果活用ステージにおける情報通信技術の活用</b>		<b>65</b>
1	調査実施ステージにおける示唆・課題	66
	1) 情報通信技術の位置付け	
	2) 情報通信技術の活用目的	
	3) 統計調査の回答者、回答内容を踏まえた活用	
	4) 検討の体制	
2	結果活用ステージにおける示唆・課題	70
	1) 情報通信技術を活用した情報提供内容	
	2) 情報通信技術を活用した情報の提供方法	
<b>附属資料目次</b>		<b>71</b>
I	国内における情報通信技術の活用に関する取組	
II	海外における情報通信技術の活用に関する取組	
III	参考ウェブサイト・文献一覧	

## **第1章 調査研究と結果の概要**

### **1. 調査研究の目的**

近年の情報通信技術の急速な発展や、ビッグデータの活用に対する関心の高まり等を受けて、「経済財政運営と改革の基本方針」(平成 25 年6月 14 日 閣議決定)を始めとする各種の政策方針において、統計データのオープン化、オンライン調査等の推進が示されるようになってきている。

「公的統計の整備に関する基本的な計画」(平成 21 年3月 13 日閣議決定)においても、公的統計が直面する課題として情報通信技術の発展に伴う統計利用ニーズの多様化が挙げられている一方で、政府統計共同利用システム(e-Stat、オンライン調査等)の活用等による府省間でのデータ共有や提供の推進についても挙げられている。

さらに、例えばアメリカセンサス局ではスマートフォンやタブレット端末用の統計データ閲覧用アプリケーションを開発し一般に提供するなど、諸外国でも統計データに対する情報通信技術の活用が進んできている。

そこで本調査研究では、今後の統計委員会における各種審議及び統計調査実施部局における検討に資することを目的として、統計の作成、提供等における情報通信技術の活用状況に関して、国内外の統計作成機関の先駆的な取組事例等を収集・整理を行った。合わせて、情報通信技術を今後活用する際の示唆・課題についても整理を行った。

# 第1章 調査研究と結果の概要

## 2. 調査研究の方法

### 1) 調査研究の全体フロー

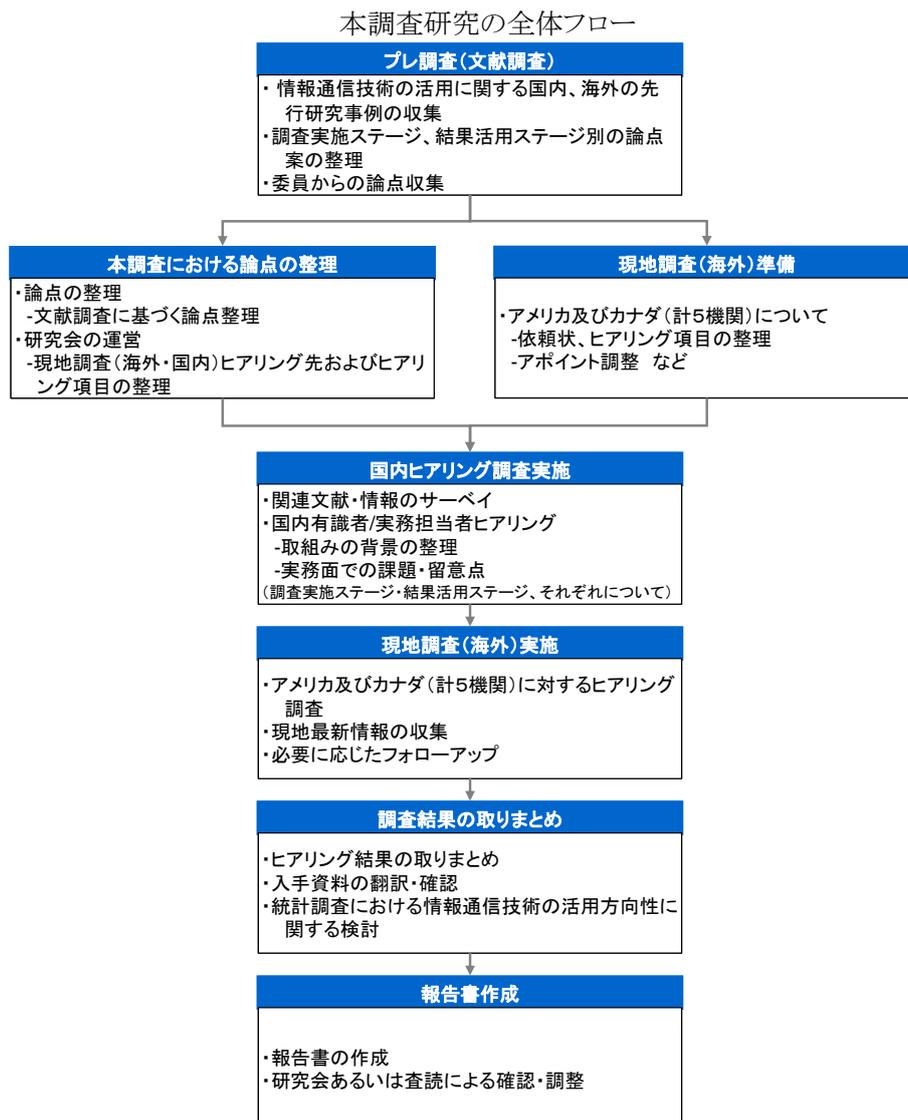
本調査研究では、まずプレ調査として、文献・書籍、各種学術機関(大学、学会等)、及び統計機関のウェブサイト等から得られる情報を収集・整理した。プレ調査(文献調査)では国内及び海外の調査における情報通信技術の活用状況について事例を収集、整理することで、本調査研究における論点整理、及び委員からの事前の論点収集を行った。

その上で、有識者研究会を実施し、本調査研究においてフォーカスすべき論点及び調査事項の整理を行った。

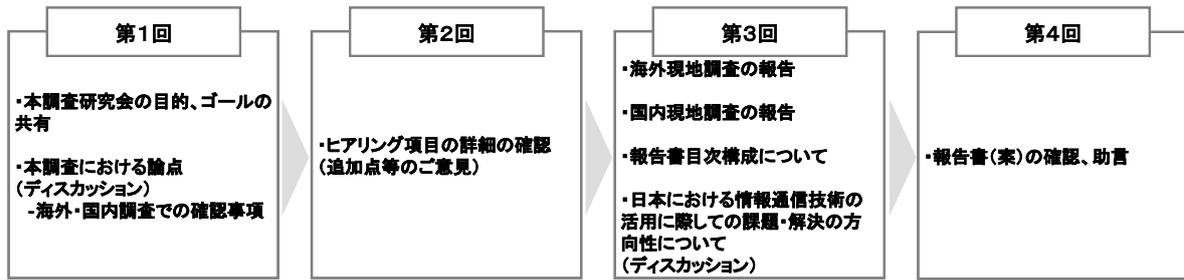
有識者研究会での検討を通じて、現地調査(海外)でなければ把握が難しい事項を中心として、現地調査による情報の深掘りを行った。海外における取組事例については、1950年代からコンピューターを活用した統計調査を実施し、情報通信技術の活用についての研究が活発なアメリカ、カナダを対象とすることとした。

また、国内における取組事例については、実務担当者(総務省統計局及び統計センター)に対するヒアリングを実施することで、実務的な課題・留意点を意識した検討を行うことができるよう配慮した。

以上を踏まえ、研究会において「統計の作成、提供等において情報通信技術を活用する際の示唆・課題」についてディスカッションを行い、その結果を最終的な報告書として取りまとめた。



本調査研究における研究会の位置付け



本調査研究において調査対象とした調査(国内)及び調査機関(海外)

<国内>

総務省統計局  
統計センター

国勢調査(平成22年国勢調査・平成27年国勢調査)  
国勢調査(平成22年国勢調査・平成27年国勢調査)

<海外>

アメリカ

米国統計局(US Census Bureau)  
米国労働統計局(BLS: US Bureau of Labor Statistics)  
ミシガン大学(University of Michigan) ICPSR  
シカゴ大学(University of Chicago) NORC

カナダ

カナダ統計局(Statistics Canada)

本調査研究においては、後述するように統計調査の実施ステージと結果活用ステージに分けた整理・検討を行うことから、公的統計作成機関に限らず、大学等を中心とする研究機関も調査対象とした。

2) 有識者による研究会の設置

本調査研究では、文献調査、国内における研究事例、海外事例調査等の情報を基に、有識者の知識・視点等を反映するため、学識経験者を中心とする4名から構成される研究会を設置し、適宜、助言を受けながら研究を進めた。

「統計の作成、提供等における情報通信技術の活用に関する調査研究」  
研究会委員名簿

座長	岩井 紀子	大阪商業大学 JGSS 研究センター長
	伊藤 伸介	明海大学経済学部准教授
	佐藤 慶一	専修大学ネットワーク情報学部准教授
	前田 忠彦	統計数理研究所データ科学研究系准教授

なお、海外現地調査には、研究会座長の岩井紀子が同行し、具体的な取組内容に関する質問や学術面からのアドバイス・支援を受けた。

## 第1章 調査研究と結果の概要

### 研究会における検討経過

回	実施時期	アジェンダ
第1回	2013年 11月19日(火)	<ul style="list-style-type: none"> <li>本調査研究の目的・概要等について</li> <li>本調査における論点(ディスカッション)               <ul style="list-style-type: none"> <li>-海外・国内調査での確認事項(対象調査・訪問先機関等のご推薦・ご紹介のお願い)</li> </ul> </li> <li>※調査実施ステージ、結果活用ステージそれぞれについて</li> </ul>
第2回	2013年 11月下旬	<ul style="list-style-type: none"> <li>ヒアリング項目の詳細の確認</li> </ul>
第3回	2014年 2月24日(月)	<ul style="list-style-type: none"> <li>海外現地調査の報告</li> <li>国内現地調査の報告</li> <li>※調査実施ステージ、結果活用ステージそれぞれについて</li> <li>報告書目次構成について</li> <li>日本における情報通信技術の活用にあたっての課題・解決の方向性について(ディスカッション)</li> </ul>
第4回	2014年 3月上旬	<ul style="list-style-type: none"> <li>報告書(案)の確認、助言</li> </ul>

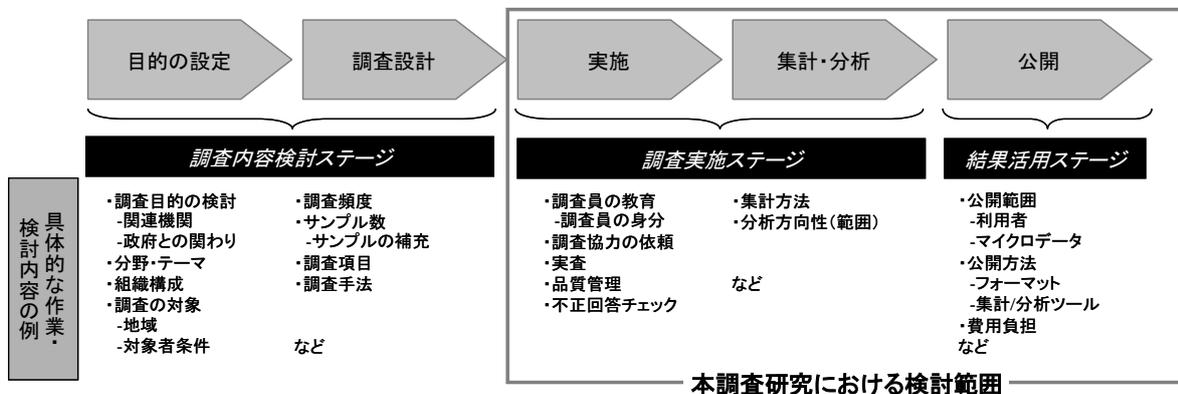
注) 会議形式での委員会開催に加え、郵送・メール及び電話による意見収集により実施した。

### 3) 国内調査・海外調査の視点

国内においては、情報通信技術を活用した取組に試行段階のものが多いことを踏まえ、アメリカ・カナダにおける統計調査において、どのような活用が行われているかを中心に整理を行うこととした。

整理にあたっては、実務への適用を意識し、統計調査の実施ステージごとの整理を行った。

### 本調査研究における整理・検討の視点



統計調査の実施ステージごとに分けた検討を行うにあたり、情報通信技術の活用による効果が高いと考えられる「調査実施ステージ」及び「結果活用ステージ」にフォーカスすることとした。特に「調査実施ステージ」では、実査時のPC活用といった調査手法における情報通信技術の活用について、「結果活用ステージ」ではデータ分析に際しての情報通信技術の活用について、検討を行うこととした。

なお、海外調査の訪問先及び主なヒアリング項目は次ページのとおりである。

## 海外現地調査訪問先リスト

国名	都市・日付	訪問先	ヒアリング項目	ステージ
アメリカ	ワシントン DC 12/9(月)	米国労働統計局 US Bureau of Labor Statistics (BLS)	・CES(雇用状況調査)における取組	調査実施
	ワシントン DC 12/11(水)	シカゴ大学 NORC University of Chicago NORC	・data enclave における取組	結果活用
	アナーバー 12/12(木)	ミシガン大学 ICPSR University of Michigan	・ICPSR における取組	結果活用
	ニューヨーク 12/13(金)	米国統計局 US Census Bureau	・Centurion(インターネットによるデータ収集システム)について ※12/10(火)大雪により US Census Bureau が閉鎖。TV 会議による実施。	調査実施
カナダ	オタワ 12/16(月)	カナダ統計局 Statistics Canada	・LFS(労働力調査)における取組 ・カナダ統計局における統計データ利活用に関する取組	調査実施 結果活用

## 第1章 調査研究と結果の概要

### 3. 調査結果の概要

#### 1) 調査実施ステージ

日本国内において情報通信技術を活用した取組を行っている例として、国勢調査が挙げられる。平成 22 年度国勢調査では、東京都をモデル地域とするインターネット回答を実施しており、平成 27 年度国勢調査に向けてさらなる検討が行われている。

一方、海外においては、複数の回答方法の1つとしてインターネット回答や PC を利用した調査員調査等が行われている。

次ページ以降において、日本、アメリカ及びカナダを対象として、調査実施ステージにおいて情報通信技術を活用した取組を行っている統計調査の概要について整理を行っている。

### 2) 結果活用ステージ

結果活用ステージにおいては、日本国内では e-Stat を総務省統計局が提供している。政府統計の総合ポータルとして機能しており、集計表や GIS 表示ツール等が提供されている。

海外においては、学術研究機関を中心に、マイクロデータの提供を含めた情報提供が行われている。「統計データの二次利用促進に関する研究会 報告書」(平成 20 年 10 月統計データの二次利用促進に関する研究会)においても、統計データの二次利用について、海外事例等の整理が行われている。

次ページ以降では、日本に加えて、アメリカ、カナダ、欧州及びスウェーデンにおける取組の概要を整理している。

# 第1章 調査研究と結果の概要

(調査実施ステージ 概要一覧 1/2)

国名	日本	アメリカ	
統計調査名(英語)	Census	Census	American Community Survey (ACS)
(日本語)	国勢調査	国勢調査	アメリカ地域社会調査
実施主体	総務省	US Census Bureau	US Census Bureau
調査対象	世帯	世帯	世帯・個人
調査目的	・人口・世帯の実態を明らかにする	・下院議会の定数を決定するとともに、連邦政府から州政府への予算配分決定のために利用される	・10年ごとに実施される国勢調査の間を埋める形で、各自治体における政策・予算配分に関する意思決定を行うために利用される
実施頻度	5年毎	10年毎	毎年
サンプル数	日本国内に住んでいる全ての人、世帯(2005年調査では回収率95.6%)	1億3,170万4,730住戸への郵送	世帯票:約327万世帯/回収率65.0% 集合住宅調査:約20万人/回収率72.6% (2011年調査)
調査手法	・国勢調査員が各世帯に調査票と説明資料などの一式を配布して、調査票の記入と提出の依頼 ・面会が不可能な場合は聞き取り調査を実施 ・紙媒体への記入後、郵送または訪問調査員に提出	・紙調査票による郵送自記調査 ・回答は、郵送による返送に加えて、調査員の訪問によるフォローが行われる	・紙調査票による郵送自記調査 ・回答は、郵送による返送に加えて、電話、調査員の訪問、ウェブでの回答が可能となっている
調査手法の変更経緯 (時期・理由など)	・自宅PCによるオンライン記入を試験導入(H22年調査、東京都全域) ・モバイル端末による記入も導入予定(H27年調査)	・2000年調査では、Long-Formと呼ばれる詳細な調査票を約300万世帯に配布していたが、2010年はShort-Formと呼ばれる簡易な調査票のみ(右記、ACS参照) ・2010年調査では、インターネットによる回答方式を導入しようとしたが、開発コストやシステム負荷の問題から断念	・2000年国勢調査で用いられた、回答項目数の多い“Long Form”調査票を、2010年国勢調査では廃止し、ACSに移行

## 第1章 調査研究と結果の概要

(調査実施ステージ 概要一覧 2/2)

国名	アメリカ	カナダ
統計調査名(英語)	Current Employment Survey (CES)	Labour Force Survey (LFS)
(日本語)	雇用状況調査	労働力調査
実施主体	US Bureau of Labor Statistics	Statistics Canada
調査対象	事業所・企業	世帯・個人
調査目的	・12の賃金区分に基づく各事業所における従業員数を回答	・雇用・非雇用状況の把握を行う ・生産年齢にあたる人口を「雇用状態」「非雇用状態」「非労働力人口」の3つに区分することで、労働力市場の状態を把握
実施頻度	半年毎(5月・11月)	毎月
サンプル数	回答率 77.3% (1,110,296 事業所のうち、858,474 事業所が回答) (2011年5月調査)	回答率 95.1% (2005年平均)
調査手法	・郵送による自記調査が主 ・電話、電子メール、調査員による訪問での回答も行われている	・CATI及び調査員による訪問面接(ノートPCを用いたCATI) ・電話番号が不明の場合には、初回は訪問した上で、残る5か月間についてはCATIを用いる
調査手法の変更経緯 (時期・理由など)	・2012年調査では、郵送による自記調査に加え、オンラインでの回答も可能になっている	・調査員の移動負荷軽減のため、2009年の調査からCAPIを導入。同時期のStatistics Canadaが実施しているCCHSやSHSでも同様にCAPIでの調査を導入している

第1章 調査研究と結果の概要

(結果活用ステージ 概要一覧 1/2)

国名	日本	アメリカ		
アーカイブ/組織名	e-stat	ICPSR	NORC	
運営主体	総務省統計局	ミシガン大学	シカゴ大学	
運営目的	各府省の統計データを1つにまとめ、さらにニーズの高い提供機能を備えることで、より利便性の高い統計データを提供	社会科学に関する調査の個票データを世界各国や国際組織から収集、保存し、それらを学術目的での二次分析のために提供	社会科学調査のデータを提供し、意思決定を支援する	
	有償/無償	原則として無償	無償	無償
	マイクロデータ提供	あり	あり	あり
	ツール提供	集計表作成ツール、GIS表示ツールを提供	秘匿管理、個人情報保護、集計、データ配布ツールを提供	シンクライアント環境によるデータアクセスが可能
	その他	利用者の要望に応じたオーダーメイド集計を実施(有償)	個票データの単純集計、クロス集計、回帰分析等、個票データのハンドリングが可能	“NORC data enclave”において、メタデータの文書化、データ分析サービス、GIS データを提供
収録データ分野	全分野、国内公的統計調査が主	政治、社会、歴史、経済、高齢者、犯罪、公衆衛生、教育、法律及び国際関係など	総合的社会調査	
利用方法・利用範囲	<ul style="list-style-type: none"> <li>・申し込みなしで利用可能(オーダーメイド集計を除く)</li> <li>・利用者の PC 上での操作が可能(オンライン接続)</li> </ul>	<ul style="list-style-type: none"> <li>・登録が必要</li> <li>・利用者の PC 上での操作が可能(一部機能はソフトウェアをインストールする必要あり)</li> </ul>	<ul style="list-style-type: none"> <li>・登録が必要</li> <li>・必要に応じて NORC スタッフによる分析支援を受けることが可能</li> </ul>	

(結果活用ステージ 概要一覧 2/2)

国名	カナダ	欧州	スウェーデン
アーカイブ/組織名	Research Data Centres (RDC)	eurostat	MONA
運営主体	Statistic Canada	European Comission	Statistic Sweden
運営目的	カナダ国内の社会情勢についての展望や、規模の異なる国の間での社会科学研究的の機会を提供する	EUの意思決定、EUの政策評価のためのデータ提供	研究者によるデータ分析を可能にするために、マイクロデータを分析するプラットフォームを提供
	有償/無償	無償	無償
	マイクロデータ提供	あり	あり
	ツール提供	ブートストラップ法による分散推定ツールを提供	集計表作成ツールを提供
	その他	GISデータの提供	データの取得はできないが、オンライン上にあるSQL形式でデータベース化された情報を、統計的に分析することが可能(クロス集計、回帰分析、因子分析など)
収録データ分野	LFS、SHS、CCFSなど、カナダ国内における統計情報	全般的・地域統計、経済・財政、人口・社会動態、産業・商業・サービス、農業・漁業、対外貿易、運輸、環境・エネルギー、科学・技術	スウェーデン国内における公的統計調査が主
利用方法・利用範囲	・原則、CRDCNに属する機関の研究者に対しては、無償でデータを提供 ・マイクロデータの利用には、申請と利用許諾契約が必要	・会員登録が必要	・利用には審査が必要 ・国外でのデータ利用についても、秘密保護法によって制限あり。基本的に禁止。

## 第1章 調査研究と結果の概要

## 第2章 統計調査において活用される情報通信技術の概要

本調査研究においては、統計調査において活用される情報通信技術について調査を行うため、「調査実施ステージ」及び「結果活用ステージ」のそれぞれについて、どのような情報通信技術が活用されているのかについて整理を行うこととした。

### 1. 調査実施ステージ

調査実施ステージにおいては、調査における回答データ収集に用いられる情報通信技術について整理を行った。まず、データ収集の方法については、大きく分けると回答者自身が回答を記録する「自記入式調査」と調査員が記録する「調査員調査」に分類される。実際の調査においては、調査票の配布方法（郵送、面接員による配付など）、調査票の形態（紙、PC など）、回答の記録者（調査員あるいは回答者）といった組み合わせに応じて、「郵送・留置調査」「面接調査（CAPI 利用）」などの分類を行う場合がある。

- ・自記式留置調査
- ・自記式郵送調査
- ・面接調査
- ・RDD 方式の電話調査（抽出の観点から、方法はCATI）
- ・CAPI: Computer Assisted Personal Interview
- ・CASI: Computer Assisted Self-administered Interview（面接調査の中で用いる場合と、インターネットネット回答など、対象者単独での回答と両方を含む）

CATI: Computer Assisted Telephone Interview

本章では、データ収集方法の分類に基づいて、それぞれのメリット、デメリットについても整理を行った。

#### 1) データ収集方法の分類

##### (1) 自記入式調査

自記入式調査においては、回答者は調査員による支援を受けずに、自分自身で調査票に回答を記録する。回答者への調査票配布（チャンネル）は、郵送やファックスといったものが挙げられる。インターネット回答も自記入式調査の1分類とみなすことが可能である。PC 等を用いる場合は CASI(Computer-Assisted Self Interview: コンピューター支援による自記入式調査)、と呼ばれる。自記入式調査を行う場合には、回答者にとって明確で分かりやすい調査票であることが重要である。

自記入式調査のメリット、デメリットとしては以下のような点が挙げられる。

<p>メリット</p> <ul style="list-style-type: none"> <li>-調査員調査と比較すると<u>調査費用を抑えることが可能</u>となる</li> <li>-<u>大規模サンプル</u>の調査に向く</li> <li>-調査員が介在しないため、<u>センシティブな内容を確認する調査</u>に向く</li> </ul>
<p>デメリット</p> <ul style="list-style-type: none"> <li>-<u>回答率が低くなりがち</u>で、督促作業等の<u>フォローアップが欠かせない</u></li> <li>-回答エラーなどが発生しやすく、<u>データ品質を高めることが難しい</u></li> <li>-郵送による配付・回収を行う場合は<u>調査期間を長く設定する必要がある</u></li> </ul>

## 第2章 統計調査において活用される情報通信技術の概要

### (2) 調査員調査

前述の自記入式調査と比較すると、調査員調査では調査員が介在するため、回答エラーの防止やデータ品質の向上を図ることが可能となる。特に、調査テーマや調査票の内容が複雑である場合、調査員調査が有効な手法となり得る。調査員調査のメリット、デメリットは以下のように整理される。

メリット -調査員を通じて、調査に対する回答者の関心を高めるなど回答率を高める取組を行いやすい -回答エラーの発生を防ぐことができる(非回答の発生防止など) -回答内容の確認など、実査後のフォローアップ作業を減らすことができる -必要に応じて、調査員を増員して回収作業を早めることも可能となる
デメリット -調査員に対する人件費や旅費等のコストが必要となる -調査員に対するトレーニングが必要となる(熟練した調査員が求められる)

さらに、調査員調査は実際に調査員と回答者が対面して行う「面接調査」と、電話によるやり取りを行う「電話調査」に分類される。

#### ・面接調査

調査員調査では、回答者による調査票の記入に対して調査員による支援が行われる。調査は面接で行われることが多く、回答者の住居や職場など様々な場所で行われる。特に、紙による調査票を用いる場合は PAPI(Paper And Pencil Interview: 紙と鉛筆による調査)、PC 等を用いる場合は CAPI(Computer-Assisted Personal Interview: コンピューター支援による面接調査)、と呼ばれる。

#### ・電話調査

調査員が介在する形の調査の中でも、電話を利用して回答データを収集する場合がある。紙による調査票を用いる場合もあるものの、PC 等を利用する場合を特に CATI(Computer-Assisted Telephone Interview: コンピューター支援による電話調査)と呼ぶ。

CATI を利用することで、前述のとおり、回答エラーの発生を防ぎ、データ品質の向上を図ることが可能となる。また、調査員を僻地へ派遣する必要もなくなる。一方で、電話調査では以下のような点に留意する必要がある。

図表1 電話調査における留意点

-提示物(イラストや写真など)を用いることができない - <u>複数回答形式の質問は、単数回答形式の質問に分解しないと回答しにくい</u> -同音異義語を避けなければならない -電話を受ける場所に他人がいることも想定されるため、繊細な内容を質問しにくい -面接調査に比べると、調査票のボリュームを抑える必要がある
--

## 2) コンピューター支援によるデータ回収 (CAI: Computer-Assisted Interview)

調査実施ステージにおけるデータ回収にあたってコンピューターを活用するものを CAI(Computer-Assisted Interview)と呼ぶ。具体的には、以下に挙げる3つの方法が代表的なものである。

- ①CASI(Computer-Assisted Self Interview: コンピューター支援による自記入式調査)
- ②CATI(Computer-Assisted Telephone Interview: コンピューター支援による電話調査)
- ③CAPI(Computer-Assisted Personal Interview: コンピューター支援による面接調査)

一般的に、PAPI に比べて CAI はデータ品質の向上やデータ管理をより容易に行うことができる。例えば、質問間の整合性や、場合によっては、過去の回答内容との整合性を、実査の時点で確認することができる。また、実査の進捗状況管理(回収数のモニタリングなど)もより容易になる。調査票上では、プログラムを組み合わせることで、より複雑な構造の調査票でも、回答エラーを防ぐことが可能となる。

一方で、調査員あるいは回答者が回答を記録するにあたって、ソフトウェアに一定程度の習熟が必要とされる。加えて、ソフトウェア上で構築されたプログラムが確実に動作するかを、事前に入念に確認する作業も必要とされる。

### (1) CASI(Computer-Assisted Self Interview: コンピューター支援による自記入式調査)

CASI による調査は、回答者にとって都合のよい時に回答できる一方で、回答者自身が PC やソフトウェアを保有している必要がある。近年では、インターネットを利用した CASI が普及している。

インターネット回答では、回答者は自身の PC を利用して、オンライン上の調査票に回答を行うこととなる。インターネット回答のメリットとしては、データ回収のコストを下げることが可能となること、回収時間の短縮化が可能となることが挙げられる。一方で、回答者自身の PC を利用することから、異なるバージョンの OS(オペレーティングシステム)やブラウザにも対応できること、暗号化通信を行うなど情報のやり取りにおけるセキュリティ確保が必要となる。

### (2) CATI(Computer-Assisted Telephone Interview: コンピューター支援による電話調査)

CATI では、調査員が PC の画面上に表示される調査画面に沿って、読み上げを行い、回答を記録していく。「図表1 電話調査における留意点」と同様の内容に留意する必要があるが、紙による調査票に代わって PC を利用することで、より複雑な調査票に基づいた調査を行うことができる、データ入力の手間をなくすことで調査期間の短縮化を図ることができる、といったメリットが生まれる。

### (3) CAPI(Computer-Assisted Personal Interview: コンピューター支援による面接調査)

CAPI では、調査員による面接調査において、PC を用いながら質問、回答記録を行う。調査員がそれぞれ PC (ノート PC やタブレット PC など) を所持し、その PC 上に表示される調査画面に従って調査を行い、回答を記録する。

CATI や CAPI で用いられるソフトウェアには、オランダ統計局が開発、提供している blaise、ノルウェーの民間企業が開発している confirmit などが存在する。また個別の調査機関が独自に開発を行うケースもあるが、基本的な機能として調査票のスキップパターンの制御やデータ整合性チェックなどを行うことができるものである。

CASI で利用されるハードウェア、ソフトウェアの費用については、複数の調査で共有することで、1調査あたりの負担を減らすなどの取組も必要とされる。特に、新規に CAI を導入する場合には、プログラミングや検証に係る時間も必要とされる。

## 第2章 統計調査において活用される情報通信技術の概要

図表2 CAI のメリット・デメリット

	メリット	デメリット
CAI 全般	<ul style="list-style-type: none"> <li>・回答時にエラーチェックを行うことができるため、実査終了後の再確認の手間を省ける</li> <li>・データ収集、修正、集計が統合されるため、紙による調査に比べて、時間を節約できる</li> <li>・データ精度が向上する可能性がある</li> <li>・複雑な構造の調査票を構成できる</li> <li>・回答率のモニタリングなど、品質管理が容易である</li> <li>・繰り返し実施される、あるいは大規模の調査であるほど、紙による調査よりも低コストで実施できる</li> <li>・類似の調査で用いられているプログラムを流用することでコストを低減させる可能性がある</li> <li>・完了票の秘匿管理が容易である</li> <li>・紙資源の節約につながる</li> </ul>	<ul style="list-style-type: none"> <li>・ソフトウェアの開発に時間及び費用が必要となる場合がある</li> <li>・調査全体を通じて、プログラムが確実に動作するかをテストする必要がある</li> <li>・プログラム開発や実査において、技術に精通した人材を確保する必要がある</li> <li>・技術的な問題が発生する可能性がある</li> <li>・調査員及び回答者がソフトウェアを適切に扱うことができる必要がある</li> <li>・ノートPC やタブレットPC などのハードウェアが必要となる</li> <li>・実査に用いた PC からデータサーバー等へのデータ移行を安全に行える環境が必要となる</li> </ul>
CASI	<ul style="list-style-type: none"> <li>・回答者にとって、回答しやすい (インターネット経由での回答も同様)</li> </ul>	<ul style="list-style-type: none"> <li>・回答者が PC を扱える必要があり、回答に用いられるソフトウェア、ハードウェアを使いこなせる必要がある</li> <li>・(インターネット回答の場合) 本人確認が困難である</li> </ul>
CATI/CAPI	<ul style="list-style-type: none"> <li>・紙による調査に比べると、訪問日時の管理といった、調査全体の計画を管理しやすい</li> <li>・PC 操作に慣れた調査員であれば、回答者との応対に注力できるため、回答データの品質が向上する可能性がある</li> </ul>	<ul style="list-style-type: none"> <li>・調査員に対する教育が必要となる</li> <li>・PC の盗難など、安全対策を講ずる必要がある</li> </ul>

### 3) 調査実施ステージにおいて活用される情報通信技術

ここでは、CAI で活用される技術の内容について整理を行う。

#### (1) インターネット

インターネットを通じた調査実施については、web サイト上での回答や e メールによる回答などが該当する。近年のインターネットの普及に伴って、特に web サイトへアクセスし回答を行う形での調査も実施されている。回答者側では、PC (スマートフォンやタブレット PC を含む) などのハードウェアとブラウザ、インターネット回線を用意する必要があるものの、必要に応じて、動画や音声、画像情報などの提示もできるなど、より多くの情報を得ることができる重要な技術といえる。

#### (2) TDE(Touchtone Data Entry:タッチトーン入力)

TDE(Touchtone Data Entry: タッチトーン入力)とは、電話調査において、回答者が電話のトーン発信を利用して回答するものである。コールセンターへの架電にあたって、会員番号の入力などを行う際に用いられる技術を、調査に利用している。回答者は電話機のキーパッドを利用して、例えば、「はい」ならば「1」を、「いいえ」ならば「2」をダイヤルする、といった形で回答を行う。

TDE は、その回答方法から以下のような特性を持った調査に向いている。

- 調査票内容が短い
- 調査内容が繰り返し実施される
- 数量値を回答する

#### (3) VR(Voice Recognition: 音声認識)

VR(Voice Recognition: 音声認識)とは、回答者が発言した内容をコンピューターが聞き取った上で記録するものである。VR 自体は、様々な分野での利用がなされており、携帯電話の音声入力機能などにも用いられている。前述の TDE と比較すると、タッチトーン式の電話を保有していない回答者でも利用できることがメリットとなる。

一方で、音声認識技術自体の精度にも影響される。コンピューターが認識した回答内容が正しいか否かを、回答者に再度確認する必要があるなど、回答時間が長くなりがちとなる欠点も挙げられる。

#### (4) EDI(Electronic Data Interchange: 電子的データ交換)

EDI(Electronic Data Interchange: 電子的データ交換)とは、財務情報等のデータを、システムを通じて一定の様式で抽出、送信するものである。いったん様式が確定すれば、繰り返し行われる調査においてデータの回収コストを大きく削減することが可能となる。必要に応じて専用線を利用してセキュリティ確保を図る場合もある。

しかしながら、データ様式の調整やシステム構成の構築など、セットアップに係る時間、費用がかかるため、EDI を利用することができるのは、企業規模が大きい回答者に限定されるといった制限が存在する。

#### (5) その他

上記以外の情報通信技術としては、海外の家計調査等で活用が検討されているレシートスキャン(画像認識・読取り技術)やバーコードスキャンといったものが挙げられる。読取り精度に起因する課題や多岐にわたる様式への対応(レシートのフォーマットやバーコード情報の更新、追加など)といった課題が指摘されるものの、統計調査への活用について海外を中心として研究・検討が行われている。

## 第2章 統計調査において活用される情報通信技術の概要

### 2. 結果活用ステージ

結果活用ステージにおいては、回答データを集計、分析する際に用いられる情報通信技術について整理を行った。分析前のデータ作成を含めて、比較的単純な集計表の作成や視覚化などを「データ作成・集計機能」として、相関分析や回帰分析等が求められるものを「分析機能」として、それぞれにおいて活用される情報通信技術を整理している。

#### 1) データ作成・集計機能

ここでは、基本統計量の算出を行う集計表の作成や、集計結果の視覚化に用いられる情報通信技術の概要を整理する。なお、下記2つの技術以外にも、e-stat において提供されているクロス集計を行う機能を実現するためのデータベース技術が存在するものの、近年の技術進歩によって可能となった技術・機能について概要の整理を行った。

##### (1) コーディング技術

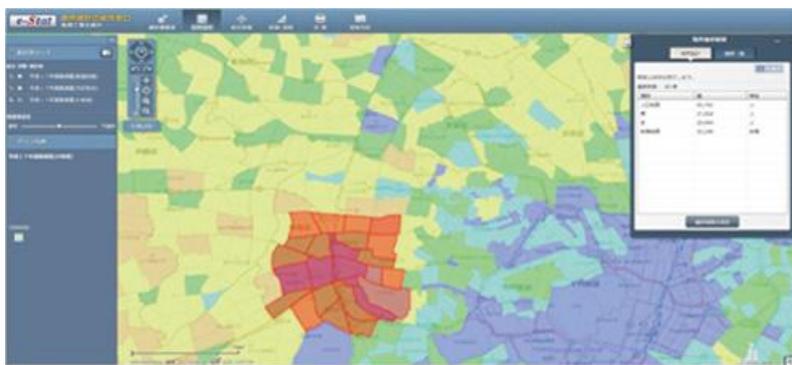
回答内容に含まれる自由回答(文章や単語での回答)を集計する際の、コーディングについても情報通信技術を活用することが可能である。従来は、手作業での確認作業が必要とされていたが、作業時間の短縮化や品質管理を高めるために、情報通信技術を活用しながら作業の自動化が図られている。

コーディング技術では、元データ(回答内容が記録されたもの)と参照ファイル(辞書の役割を果たすもの)の2ファイルを準備する。参照ファイルでは、回答内容(記述内容)とそれに対応するコード番号が管理される。元データについては、最初に表現の標準化作業が行われる。同一の内容を指す言い回しや表現を統一しつつ、意味のない記号等を削除していく。続いて、参照ファイルに基づいて、回答内容のコード化が行われる。上記作業を行った上で、なおコード化が行われない内容については、手作業での確認が行われるものの、テキストマイニングを行うソフトウェア等を活用しながら、より効率的な作業実施が図られる。しかしながら、正しい分類が成されるか否かについて、現状では精度に問題があることから、完全な自動化は極めて困難となっている。

##### (2) GIS 技術

統計情報を地図上の表示し、集計結果を視覚化するものとして GIS 技術が挙げられる。複数の情報を同一地図上に表示するなどの機能も実現が可能であり、地理的な情報を分析することができる。

(参考)総務省統計局 e-stat における「地図で見る統計(統計 GIS)」より



### 2) 分析機能

集計に加えて、より複雑な分析を行うにあたって活用される情報通信技術について、以下のとおり整理を行った。相関分析や回帰分析など行うために、マイクロデータに対してアクセスすることを前提とすると、以下の2つの技術が重要となる。これらはいずれもデータ保護を行う上でも重要な技術である。

#### (1) リモートアクセス技術<sup>1</sup>

リモートアクセスとは、マイクロデータを保持するサーバーに、データ利用者がアクセスして分析を行うことを指す。マイクロデータをサーバーから持ち出すことなく、異なる切り口での分析を試行することができるため、統計データのより高度な利活用につながる技術である。アクセスにあたっては、専用の施設を設置する場合に加えて、専用端末の設置、個人 PC からのアクセスなどが行われる。

基本的に、元のマイクロデータ自体を参照することはできず、利用者は分析結果のみを得ることとなる。そのために、データ持ち出し防止や、個人が特定されうる分析結果の出力抑制といった機能も備えられる。

#### (2) シンククライアント技術

シンククライアント技術とは、データ及びアウトプットの保存機能を持たない端末を利用して、サーバー上のプログラムを実行、結果を参照することを可能とする技術である。データ利用者が使用する端末は、ハードディスクなどの記憶媒体やデータを持ち出すための端末(USB 等の入出力端末)を持たず、サーバーとのやり取りを行うインターフェースとしてのみ機能する。

前述のリモートアクセスを行うにあたって、セキュリティをより高いレベルで確保するためにも、シンククライアント技術は利用されている。一般的に、サーバーはセキュリティの確保されたデータセンター等に設置される一方で、利用者が使用する端末には紛失や盗難等のリスクが存在する。シンククライアント技術を活用することで、セキュリティレベルを向上させながらも、より高度なデータ分析を行うことが可能となる。

---

<sup>1</sup>内閣府統計委員会基本計画部会第3ワーキンググループ(第2回)資料2  
「リモートアクセス」について(総務省政策統括官室提出資料)参照

## 第2章 統計調査において活用される情報通信技術の概要

### **第3章 調査実施ステージにおける情報通信技術の活用**

本調査研究では、調査実施ステージにおける情報通信技術の活用に関する取組事例として、国内及び海外の統計機関において、どのような取組がなされているかの整理を行った。

国内については、統計局および統計センターへの聞き取り調査等を行いながら、国勢調査のオンラインによる実施に関する事例を整理した。

海外については、アメリカ及びカナダにおいて、情報通信技術を調査実施ステージに活用している統計機関を対象として、取組の経緯、内容、今後の課題等に関する聞き取り調査を行った。

具体的には、アメリカセンサス局及び労働統計局、カナダ統計局の3つの統計機関における取組の整理を行うこととした。労働統計局においてはCES(Current Employment Survey:雇用状況調査)について、カナダ統計局においてはLFS(Labour Force Survey:労働力調査)について、整理を行った。

なお、聞き取り調査を行う中で、アメリカセンサス局においては、複数の統計調査に利用可能なCenturion(インターネットによるデータ収集システム)を運用していることが明らかになったため、前述のCenturionについて整理を行うこととした。

## I 国内における情報通信技術の活用に関する取組

### ・統計局および統計センター

#### ・国勢調査における取組

##### (1) 調査概要

調査名	: 国勢調査
概要	: 人口・世帯の実態を明らかにする調査
対象	: 世帯
実施時期	: 5年毎
回答方法	: 郵送または訪問調査員に提出 平成 22 年国勢調査では、東京都で自宅 PC を用いた回答方法を施行導入
抽出方法	: 全数抽出(世帯数)
回収数	: 2005 年調査における回収率は 95.1%

##### (2) 情報通信技術の位置付け

国勢調査では、回答者の利便性向上に向けた取組の一つとして、オンライン回答の全国的な導入を検討している。従来、国勢調査では、調査員により配布された調査票に回答を記入したものを調査員に直接提出する方式に加え、面会が不可能な場合には聞き取り調査が行われてきたが、平成 22 年国勢調査から、新たに郵送方式の導入及びオンライン回答方式の試行が実施された。

尚、平成 22 年度国勢調査におけるオンライン回答方式は、調査員によって各世帯に配布された調査票を基に、インターネット上で回答を実施するものであり、東京都のみでの試行導入となっている。

##### (3) 組織体制・機能

国勢調査の集計部門は、独立行政法人統計センター(以下、統計センター)となっている。統計センターでは、システムの開発・運用の一部を外注しており、リソースの制約もあることから統計センター内部で完結するようなシステム開発は限られている。

国勢調査に係る検討は、①国勢調査計画会議、②国勢統計実務検討会、③国勢調査有識者会議で実施されている(平成 25 年 4 月から、上記の体制に移行した)。

###### ①国勢調査計画会議

調査の実施部門である統計局と、集計部門である統計センター内の連絡調整・検討を行う会議であり、調査企画班、環境整備班、結果提供班、製表企画班、情報システム班の各分科会が存在している。

###### ②国勢統計実務検討会

調査方法等の実務に関して、地方公共団体の参加を得て具体的な検討を実施する会議である。

###### ③国勢調査有識者会議

体制移行前における「企画に関する検討会」と、「事例研究会」という海外事例等についての研究会を合わせたもの。国勢調査有識者会議では、国勢調査の実施に係る検討を 3 つの WG(i 企画 WG、ii 広報 WG、iii 環境 WG)で実施している。i は企画、ii は広報についての検討、iii はプライバシー意識の高まりを受け年々厳しくなっている調査環境についての検討が実施されている。

#### (4) 情報通信技術の導入経緯・目的

平成17年4月に施行された個人情報の保護に関する法律(以下、個人情報保護法)の影響により、平成17年国勢調査では、実査の段階でプライバシーを理由にして、回収に時間を要するケースが生じていた。このような状況を踏まえ、H22年国勢調査では調査の円滑な実施を目的として、情報通信技術の活用や郵送による回収といった回答方法の選択肢を増やす検討を行い、オンライン回答の導入(東京都のみ)と郵送回答の導入を実施した。

他国の事例では、コスト削減、回収効率の向上といった観点で情報通信技術を活用した回答方法が導入されていることが多いが、国勢調査では回答者の負荷軽減等、円滑な調査の実施を第一の目的としている。

#### (5) 情報通信技術の活用に関する評価

平成27年国勢調査に向け、平成24年7月(第1次)、平成25年6月(第2次)及び平成26年6月(第3次)に試験調査を実施しており、オンライン回答方式の導入推進に向けた検討を実施している。それぞれの試験調査の概要及びオンライン回答方式の導入に係る検証事項を以下に示す。

##### ・平成27年国勢調査第1次試験調査

実施期日 :平成24年7月12日(木) ※平成27年10月1日と同じ曜日を設定  
調査対象 :7府県14市町(384調査区、約19,000世帯)  
調査方式 :調査区ごとに先行方式、並行方式を分けて実施  
検証事項 :先行方式、並行方式の調査方式の違いが回答率に与える影響の検証

##### ・平成27年国勢調査第2次試験調査

実施期日 :平成25年6月20日(木)  
調査対象 :7都道府県14市区町(168調査区、約8,400世帯)  
調査方式 :先行方式  
検証事項 :若年層のインターネットの利用形態(モバイル端末等)に対応した回答方法の検証

##### ・平成27年国勢調査第3次試験調査

実施期日 :平成26年6月19日(木)  
調査対象 :都道府県庁所在市及び都道府県庁所在市以外の政令指定都市(東京都の特別区(1区)を含む52市区)の区域に属する平成22年国勢調査調査区の中から、地域特性ごとに選定する520調査区、28,080世帯  
調査方式 :先行方式  
検証事項 :調査方法、調査事務、調査設計等についての最終的な検証

### 第3章 調査実施ステージにおける情報通信技術の活用

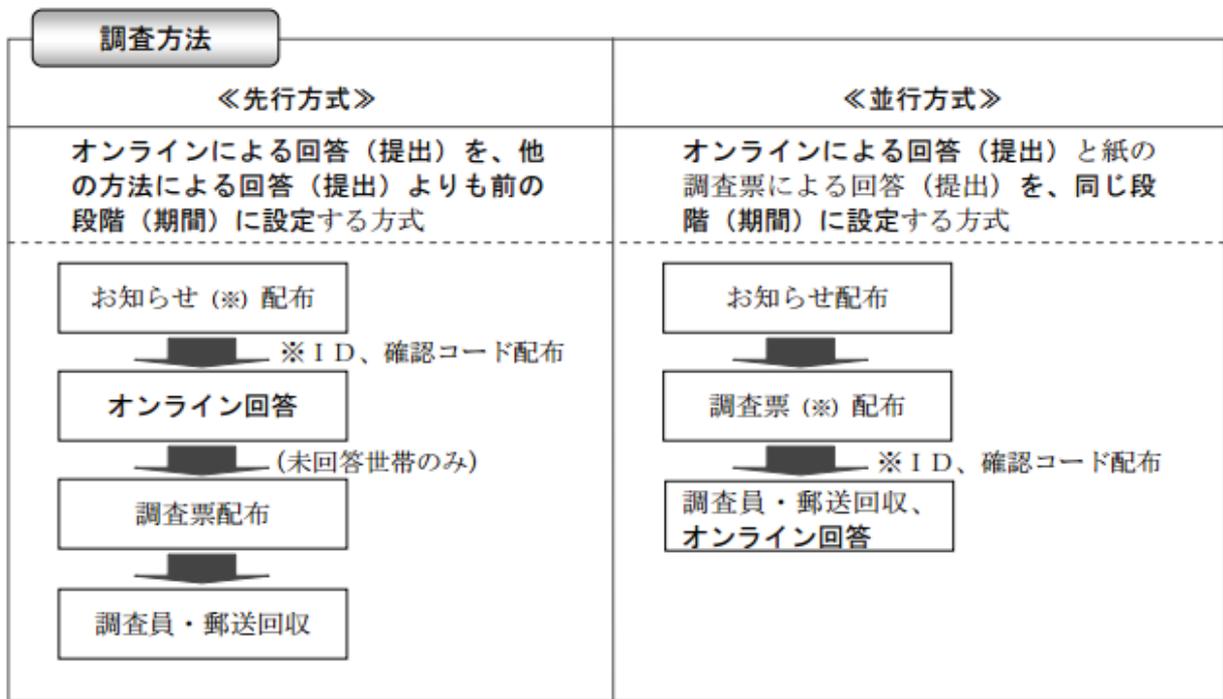
これらの試験調査の結果及び平成 22 年国勢調査の結果を踏まえ、回答率及び回答精度の観点からの情報通信技術の活用に関する評価を以下に記す。

#### ①回答率

##### i) 調査方式

平成 27 年国勢調査第 1 次試験調査では、オンラインによる回答(提出)を、他の方法による回答(提出)よりも前の段階(期間)に設定する方式である先行方式、オンラインによる回答(提出)と紙の調査票による回答(提出)を、同じ段階(期間)に設定する並行方式と、調査方式の違いが回答率に与える影響の検証がなされた。

図表1 調査方式の違い



(出所:総務省統計局「平成 27 年国勢調査第 1 次試験調査の概要」)

平成 22 年国勢調査は並行方式での実施であり、オンライン回答率は 8.3%と他国と比べて低い結果となった。平成 24 年 7 月に実施された平成 27 年国勢調査第 1 次試験調査では、先行方式と並行方式の比較がなされ、調査方式によりオンライン回答率が大きく異なる傾向が見られた(図表2)。

図表2 回答方法の割合(平成 27 年国勢調査第 1 次試験調査)

	実数(世帯)					割合(%)				
	合計	調査員	郵送	オンライン	聞き取り調査	合計	調査員	郵送	オンライン	聞き取り調査
先行方式	9,708	1,305	4,246	2,452	1,705	100,0	13,4	43,7	25,3	17,6
並行方式	9,015	2,373	4,559	584	1,499	100,0	26,3	50,6	6,5	16,6

(出所:総務省統計局「平成 27 年国勢調査第 1 次試験調査の概要」)

### 第3章 調査実施ステージにおける情報通信技術の活用

図表2を見ると、先行方式におけるオンライン回答率(25.3%)は、並行方式におけるオンライン回答率(6.5%)の約4倍となっている。また、全調査区において先行方式が用いられた平成27年国勢調査第2次試験調査の結果においても、オンライン回答率は23.3%となっており、オンライン回答の推進の観点からは先行方式を採用することが有効であると判断された。

#### ii) 世帯構成

平成27年国勢調査第1次調査及び平成27年国勢調査第2次調査の結果では、世帯人員別のオンライン回答の割合(聞き取り調査を除く)にも違いが見られた(図表3及び図表4)。

図表3 世帯人員、回収方法別 回答の割合(平成27年国勢調査第1次調査)

	合計			
	調査員	郵送	オンライン	
総数	100.0	16.3	53.0	30.7
単身世帯	100.0	14.4	63.8	21.8
世帯人員2人	100.0	18.1	51.3	30.6
世帯人員3人	100.0	15.8	48.0	36.1
世帯人員4人	100.0	14.1	46.6	39.2
世帯人員5人以上	100.0	23.5	40.6	35.9

(出所:総務省統計局「平成27年国勢調査第1次試験調査の概要」)

図表4 世帯人員、回収方法別 回答の割合(平成27年国勢調査第2次調査)

	合計			
	調査員	郵送	オンライン	
総数	100.0	37.6	32.9	29.6
単身世帯	100.0	34.0	39.3	26.7
世帯人員2人	100.0	42.3	31.4	26.2
世帯人員3人	100.0	36.7	30.3	33.0
世帯人員4人	100.0	35.5	29.4	35.1
世帯人員5人以上	100.0	41.1	22.7	36.2

(出所:総務省統計局「平成27年国勢調査第2次試験調査の概要」)

図表3及び図表4では、世帯人員別のオンライン回答の割合は、世帯人員の少ない世帯(単身世帯及び世帯人員2人の世帯)に比べ、世帯人員の多い世帯(世帯人員3人以上)で高い値となっている。この要因として、世帯人員が少ない世帯に占める高齢者の割合が高いことが考えられる。

### 第3章 調査実施ステージにおける情報通信技術の活用

#### iii) 回答に用いる端末

平成 27 年国勢調査第 1 次調査におけるオンライン回答では、PC での回答がほとんどであったが、推奨していないタブレット型端末やスマートフォンでの回答も行われていた。同調査においてオンライン回答をした世帯へのアンケートでは、約 4 割がスマートフォン等の携帯端末でのオンライン回答を希望しており、特に 39 歳以下の年齢階級では過半数が携帯端末でのオンライン回答を希望しているという結果(図表5)を得ており、携帯端末を用いた回答への対応も必要であると考えられる。

図表5 スマートフォンなどの携帯端末でのオンライン回答を希望する人の割合(%)

	総数				
		39 歳以下	40 歳～59 歳	60 歳以上	無記入・不詳
総数	100.0	100.0	100.0	100.0	100.0
希望する	40.8	57.6	38.1	19.7	39.5
希望しない	59.2	42.4	61.9	80.3	60.5

(出所:総務省統計局「平成 27 年国勢調査第 1 次試験調査の概要」)

図表5の結果を受け、平成 27 年国勢調査第 2 次試験調査では若年層のインターネットの利用形態対応した回答方法を提供する目的で、スマートフォン等のモバイル端末にも対応したシステムが構築され、PC 版、スマートフォン版と 2 つの電子調査票も用意された。同調査では、オンライン回答(23.3%)の約 8 割が PC 版、約 2 割がスマートフォン版の電子調査票を用いた回答という結果(図表6)となった。

図表6 電子調査票の種類別回答割合(%) (平成 27 年国勢調査第 2 次試験調査)

総数		
	PC 版	スマートフォン版
100.0	79.2	20.8

(出所:総務省統計局「平成 27 年国勢調査第 2 次試験調査の概要」)

#### ②回答精度

紙の調査票による回答では、単一回答の同一の調査項目に対して複数回答してしまうマルチマークや、記入漏れ等の記入ミスが生じてしまうことがあるが、オンラインによる回答では、それらの記入ミスを防止することができる。例えば、選択式の設問で回答により分岐があり、次の設問が飛んだ場合、紙の調査票による回収では記入ミスが生じやすいが、オンラインであれば誤記入、未記入を提出前にチェックすることができる。そのため、紙の調査票による回答であれば調査員による初期審査(記入漏れがないか等)と市町村による審査が必要だが、オンライン回答の場合は初期審査の負担も少なくなる。

一方で、オンライン回答に用いる調査区番号、世帯番号の入力漏れ等、オンライン回答特有の問題が生じることもある。平成 22 年国勢調査におけるオンライン回答方式(東京都でのみ試行導入)では、回答者の世帯の特定のため、回答者は ID、調査区番号に加えて世帯番号を入力することが必要であったが、世帯番号の入力漏れが多数生じるという問題が発生した。世帯番号の入力漏れが生じた調査票は「迷子票」と呼ばれており、調査区と氏名のみから世帯を特定しなければならないため、世帯の特定に要する調査員への負荷が非常に大きくなる。この問題を踏まえ、平成 27 年国勢調査では、ID と世帯番号の紐付けを行う等、「迷子票」の特定を容易にする対策を検討している。

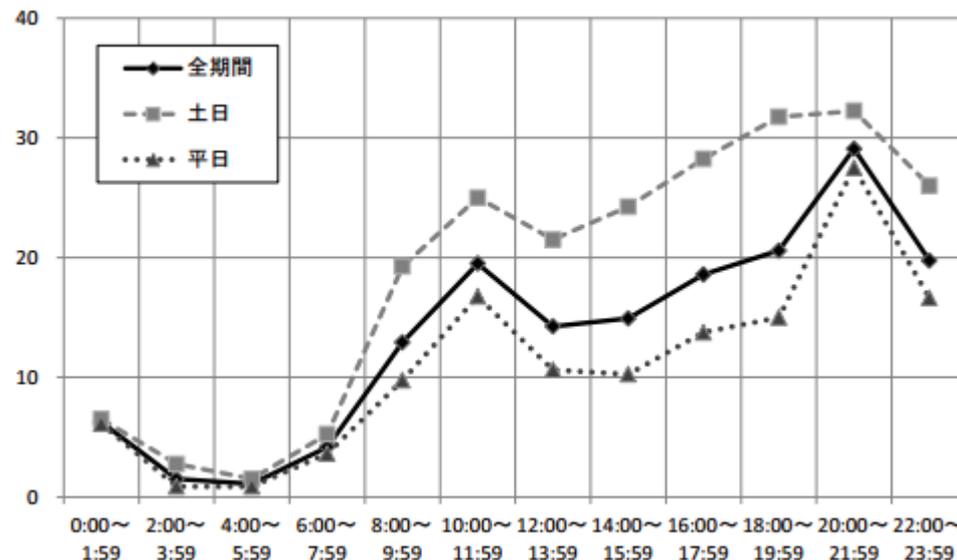
(6) 課題と今後の方向性

オンライン回答方式は、平成 27 年国勢調査において全面的に導入することが決定されており、課題と対策の方向性について、有識者会議で議論がなされている。

有識者会議において挙げられている課題の 1 つとして、ITリテラシー面でのユーザーサポートの強化がある。調査票には回答するホームページの URL が記載されており、回答にはその URL を入力する必要があるが、URL 入力に慣れていない人も多く、入力ページにたどり着けないという事例が発生している。このような ITリテラシーの面でのユーザーサポート等の対策として、URL 入力に日本語ドメインを採用する、検索サイトに直接的なリンクやバナー広告、スポンサードサーチ広告を掲示する、検索サイトにおけるキーワード検索からホームページに誘導するといった、入力ページへのアクセスにおける支援が検討されている。

また、回答時間の集中への対応も課題として挙げられている。国勢調査におけるオンライン回答は、土曜日、日曜日の夜に集中する傾向にあり(図表7)、特に日曜日は他の曜日の約2倍の回答数となることもある。このように、回答が集中する時間帯におけるサーバー負荷、コールセンターの人員体制等についても、対策が必要であり検討が進められている。

図表7 回答曜日 回答時間帯別回答数(平成 27 年国勢調査第 2 次試験調査)  
(件)



(出所:総務省統計局「平成 27 年国勢調査第 2 次試験調査の概要」)

### 第3章 調査実施ステージにおける情報通信技術の活用

尚、参考として平成 27 年国勢調査第 2 次試験調査において、オンライン回答をしなかった理由を以下に示す。

図表8 オンライン回答しなかった理由(平成 27 年国勢調査第 2 次試験調査)

オンライン回答しなかった理由	割合(%)
1. パソコンやスマートフォンを持っていないから	27.6
2. 紙の調査表の方が回答しやすいから	22.8
3. インターネットの使い方がわからないから	10.5
4. パソコンの起動が面倒だから	8.4
5. セキュリティ面に不安があるから	5.7
6.途中で止めた	4.2
7. オンライン回答するのを忘れていたから	3.9
8. オンライン回答できるのを知らなかったから	2.7
9. オンライン回答の画面に接続できなかったから	2.5
10. スマートフォンでは画面が小さいから	2.0
11. インターネット回答の利用案内(調査対象者 ID・パスワード)をなくしたから	0.6
12. 操作ガイドで推奨された利用環境になかったから	0.5
13. その他	5.7

(出所:総務省統計局「平成 27 年国勢調査第 2 次試験調査の概要」)

## II 海外における情報通信技術の活用に関する取組

### 1. アメリカセンサス局 (United States Census Bureau)

アメリカセンサス局は、商務省の下で人口・経済に関する統計情報を提供する機関である。10年に1度実施されるセンサスやACS (American Community Survey:アメリカ地域社会調査)、経済センサスといった統計調査を実施している。ここでは、複数の統計調査に利用可能な Centurion (インターネットによるデータ収集システム) について整理を行う。

#### 1) Centurion (インターネットによるデータ収集システム) を活用した取組

##### (1) Centurion の概要

Centurion は、インターネットを通じて調査設計や実査管理、調査への回答を行うことができる web ベースのプラットフォームである。センサス局では、Centurion を利用することで、回答者の回答負荷を減らすとともに、データ品質を向上させることを目指している。Centurion は以下のような特徴を持っている。

- リアルタイムでのデータ修正
- オンラインヘルプ
- 複数言語への対応
- 過去の回答内容を表示するなど動的なコンテンツ対応
- 調査票の形式で回答内容を表示

上記のような調査に関連する機能に加えて、センサス局内部において、インターン応募者の情報管理などに用いられるなど、センサス局における基盤システムとしても利用されている。

図表5 Centurion ログイン画面の例  
ACS(アメリカ地域社会調査)

The screenshot shows the login interface for the American Community Survey. At the top, it displays the 'United States Census Bureau' logo and the title 'American Community Survey'. Below this, it states 'U.S. Department of Commerce | Economic and Statistics Administration'. A welcome message reads: 'Welcome to the American Community Survey. You will need the materials we mailed to you to start the survey. All the information that you provide will remain completely confidential.' A link is provided: 'Para completar en español, oprima aquí.' The main section is titled 'Please Log In' and contains a barcode with the text 'Example User ID' pointing to a 10-digit number. Below the barcode, there is a 'User ID:' field with a hyphen and a 'Login' button. At the bottom, there is a warning box: '\*\* WARNING \*\* You have accessed a UNITED STATES GOVERNMENT computer. Use of this computer without authorization or for purposes for which authorization has not been extended is a violation of federal law and can be punished with fines or imprisonment (PUBLIC LAW 99-474). System usage may be monitored, recorded, and subject to audit. Use of this system indicates consent to monitoring and recording.' The footer includes 'OMB Number: 0607-0810', 'Approval Expires: 09/30/2014', and links for 'Accessibility', 'Privacy', and 'Security'.

図表6 Centurion ログイン画面の例  
Annual Survey of Public Employment and Payroll (公務員雇用状況調査)

OMB No.: 0007-0452 Approval Expires: 03/31/2015

U.S. DEPARTMENT OF COMMERCE  
BUREAU OF THE CENSUS

## 2013 Annual Survey of Public Employment & Payroll

SURVEY of GOVERNMENTS

Website: [Survey of Governments - Employment](#)  
Email: [govs.employ@census.gov](mailto:govs.employ@census.gov)  
Telephone: 1-800-642-4901 weekdays, 7AM to 5PM ET

**Welcome to the Annual Survey of Public Employment & Payroll**

- The due date is **April 30, 2013**.
- This is the web collection system that enables public entities to respond to the U.S. Census Bureau's Annual Survey of Public Employment & Payroll via the Internet.

**Login**

- To log in, please enter the User ID and Password indicated on your form and press the "Login" button.
- User ID and Password are case sensitive.
- If you forgot your User ID and/or Password, contact us at 1-800-832-2839 weekdays, 7AM to 5PM ET.

**Please note: You will be logged out of the system after 45 minutes of inactivity. No data will be lost.**

User ID:

Password:

**REPORT ONLINE:** It's fast and secure. Respond to this survey via the Internet at the following Web address using the supplied User ID and Password: [respond.census.gov/aep](http://respond.census.gov/aep)

User ID:

Password:

**\*\* WARNING \*\***

You have accessed a UNITED STATES GOVERNMENT computer. Use of this computer without authorization or for purposes for which authorization has not been extended is a violation of Federal law and can be punished with fines or imprisonment (PUBLIC LAW 99-474). System usage may be monitored, recorded, and subject to audit. Use of this system indicates consent to monitoring and recording.

[Burden Statement](#)   [Accessibility](#)   [Privacy](#)   [Security](#)

## (2) 開発の経緯・体制

インターネットによるデータ回収システムの開発計画は、2008 年後半から開始された。2008 年当時、5～7 つ程度の調査においては、すでにインターネットによる調査が行われていたものの、複数の調査実施を統一して管理するシステムが必要となり、Centurion が開発された。

センサス局では、外部組織の調査も実施することがあるため、Centurion の開発にあたっては、システムの柔軟性が非常に重視された。また、Centurion はセンサス局のエンジニアとプログラマーによってインハウスで開発された。当時、すでに商用化されていたソフトウェアを利用することも検討されたものの、Centurion が特に重視していた柔軟性に欠けていたため、インハウスでの開発が決定された。

基本的な枠組みは 6 カ月程度で構築され、その後、随時新たな機能が追加されていった。例えば、初期のリリース時には Centurion コントロールセンター（詳細は下記参照）は実装されていなかったが、数か月後になってその機能も追加された。

### (3) 基本的な機能

Centurin に含まれる基本的な機能は下記の 5 つである。

- 調査設計と実査(Survey Design and Implementation)
- データベース管理(Database Schemas)
- XML 形式によるデータの出力(XML Output Format)
- プロジェクト管理(Administration and Project Management)
- IT セキュリティ(IT Security)

上記の基本的な機能は、異なる調査においても共通して用いられる機能であることから、共通して利用できるフレームワークを作ること、作成時間の短縮につなげるなど、業務の効率化が図られている。

また、デスクトップ PC 向けあるいはモバイル端末(スマートフォン、ブラックベリー、iPhone など)向けのブラウザからもアクセスが可能となっている。ハードウェアとしては、デスクトップ PC に限らず、ノート PC やタブレット端末にも対応している。

### (4) 主な特徴

#### ①柔軟性の高いデザイン

Centurion では、調査内容に沿ったプログラムを記述することが可能となっており、回答者に表示される画面のデザイン(配色、レイアウトなど)の変更を行うことができる。

#### ②紙の調査票との連携

回答者は web で回答することが可能となっているが、調査票を PDF でダウンロードして印刷、参照しながら回答することもできるなど、紙の調査票による回答に慣れている回答者に対する配慮を行っている。

#### ③複数の関係者による共有

“Survey Viewer”と呼ばれるツールを用いることで、センサ局内の個別分野の専門家が、事前に調査画面を確認することも可能となっており、内容の間違いや変更修正を行うことができる。

#### ④異なる回答方法への対応

Centurion にアクセスする対象によって、画面に表示される内容が変わる(CATI の調査員に表示される調査画面と、自記入のための調査画面など)ことで、異なる回答方法にも対応可能となっている。具体的には、センサ局のコールセンターによるフォローアップ活動とも連携することが可能で、同一のシステムを利用しながら、より効率的な業務を行うことができる。Centurion によって、web による自記入式の調査画面を作成すれば、修正も必要ないか、わずかな画面修正を行うことで、CATI 調査員による質問・記入も可能となっている。

#### ⑤紙による調査票の情報を読み込んで、調査画面を自動的に生成

より効率的な業務を行うため、調査票のメタデータをもとにして、調査画面へ変換するソフトウェアも開発された。例えば、SAS(Service Annual Survey:サービス業統計)では 160 種類の調査票から、2012 年に実施された経済センサス(Economic Census)では 830 種類以上の調査票から、メタデータを抽出して調査画面への変換が行われた。これによって、開発に必要な期間が 4 カ月間程度、短縮された。経済センサスでは、30 問以上の質問があり、それぞれの質問が 75~400 項目のデータを持っているなど、大規模なメタデータであったが、この機能によって業務効率が大きく向上した。

### 第3章 調査実施ステージにおける情報通信技術の活用

また、情報通信技術を活用することで、以下のような長所も存在する。

- データ回収後の集計作業が効率化された
- 他の回答方法に比べて、データ回収のコストが削減されるとともに、データの正確性も向上した
- リアルタイムで、特定の質問項目に対する修正や検証が可能となった
- 質問ごとのヘルプ画面の設定が可能となった
- 過去の回答や前の質問に対する回答内容を表示することが可能となった
- 提出前に回答内容の確認や修正が可能となった

#### (5) セキュリティに関する取組

Centurion では、回答者からの信頼を維持すべく、情報漏洩などのセキュリティ障害に加えて、システムダウンといった障害に対する取組が極めて重視されている。

セキュリティに関しては、第三者によるセキュリティテストを行い、POAM(Process Oriented Analysis Method for Medical Incidents: 業務プロセスに着目した事故分析手法)による検証においても目立った問題はないことを継続的に確認するなど、外部機関を活用しながら対策を講じている。また、Centurion との接続はすべて https 接続を通じて行われ、安全性が確保されている。やり取りされるデータもすべて暗号化された上で、データベース化されている。

セキュリティ確保の取組の一環として、システム監査も行われている。Centurion への接続ログはすべて記録され、Centurion コントロールセンターによってモニタリングされている。ユーザーの活動履歴モニタリングも行われている。

また、回答データについても、提出後のデータ改ざんの試みを検知するための、データ整合性モニタリングが行われ、改ざん防止の取組が行われている。回答データはリアルタイムでのバックアップが作成されるとともに、FIPS 140-2 暗号化記憶媒体へのバックアップも毎晩行われている。記憶媒体は異なる場所で保管することでデータ保護の強化が図られている。

セキュリティ障害への対応に加えて、Centurion システム全体の健全性をモニタリングするためのシステムも実装されている。24 時間体制で CPU やメモリ利用率といったシステム状態が監視されており、必要な場合にはアラートが出される。

#### (6) 今後の課題

Centurion は、回答負荷の軽減に役立っているが、さらなる軽減に向けて、新たな回答ツールの開発やモニタリングのためのツール開発などが課題となっている。現在は、6つの調査において約70%の回答が、10の調査において約50%が、インターネットを通じて収集されている。Centurion による回答が選択肢として提示された対象者は約1,090万人以上に上るが、さらなる拡大が図られている。

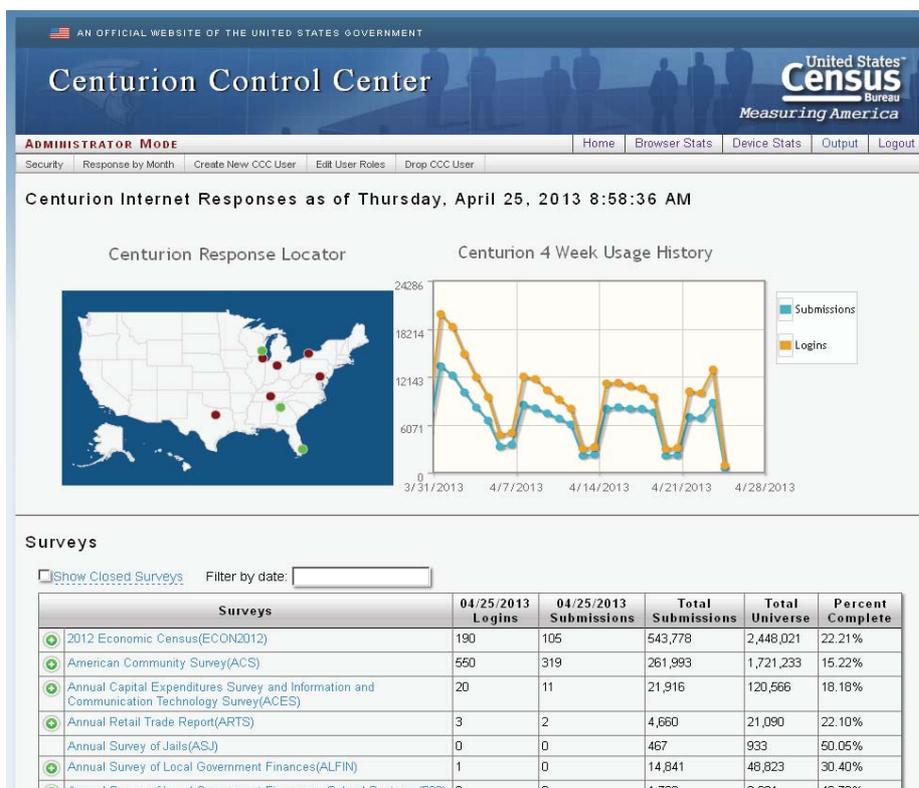
～Centurion コントロールセンター～

Centurion では、リアルタイムでのモニタリングを行うことが可能である。Centurion コントロールセンターと呼ばれるシステムでは、リアルタイムでの回収状況の把握やログ監視、システム監視等が行われる。回答数や原データの閲覧、調査票の形で回答結果を出力する機能も実装されている。主な機能としては以下のようなものがあり、回答する時間帯やどのような環境(ハードウェア、ソフトウェア)で回答が行われているかを把握することが可能となっている。

・回収状況のモニタリング

回収状況について、個別の調査に加えて、Centurion 上で行われている複数の調査を横断的にモニタリングすることもできる。

図表7 Centurion コントロールセンター(回収状況のモニタリング)



### 第3章 調査実施ステージにおける情報通信技術の活用

#### ・利用されているデバイス情報の取得

デスクトップ PC やノート PC、iPhone など、どのような端末からアクセスさせているかをモニタリングすることができる。

図表8 Centurion コントロールセンター (デバイス情報のモニタリング)

The screenshot shows the Centurion Control Center interface. At the top, it says "AN OFFICIAL WEBSITE OF THE UNITED STATES GOVERNMENT" and "United States Census Bureau Measuring America". Below that is a navigation bar with "ADMINISTRATOR MODE" and links for "Home", "Browser Stats", "Device Stats", "Output", and "Logout". A secondary bar contains "Security", "Response by Month", "Create New CCC User", "Edit User Roles", and "Drop CCC User". The main content area is titled "Centurion User Device Info as of Mar 20, 2013 12:00" and contains a table titled "Centurion Logins by Device".

Rank	Device	Logins	Percent
1	Other	1,933,578	98.04%
2	iPad	27,758	1.41%
3	Android	6,499	0.33%
4	iPhone	4,271	0.22%
5	WAP capable phone	44	0.00%
6	BlackBerry	42	0.00%
7	Sony PlayStation 3	10	0.00%
8	NOKIA Lumia 900	8	0.00%
9	NOKIA Lumia 710	8	0.00%
10	HTC Radar 4G	7	0.00%
11	SAMSUNG SGH-i677	6	0.00%
12	Amazon Kindle	6	0.00%
13	HTC T9295	5	0.00%
14	HTC HD7	4	0.00%
15	webOS	4	0.00%
16	HTC T8788	3	0.00%
17	SAMSUNG SGH-i937	2	0.00%
18	HTC P139100	2	0.00%
19	Nintendo 3DS	1	0.00%
20	HTC P186100	1	0.00%
21	SAMSUNG SGH-i817	1	0.00%

At the bottom of the interface, there is a dropdown menu labeled "Select Survey".

### 第3章 調査実施ステージにおける情報通信技術の活用

#### ・利用されているブラウザ情報の取得

インターネットエクスプローラーや firefox といった、ブラウザの種類やバージョンについてもモニタリングすることができる。

図表9 Centurion コントロールセンター (ブラウザ情報のモニタリング)

AN OFFICIAL WEBSITE OF THE UNITED STATES GOVERNMENT

Centurion Control Center

United States Census Bureau  
Measuring America

ADMINISTRATOR MODE

Home | Browser Stats | Device Stats | Output | Logout

Security | Response by Month | Create New CCC User | Edit User Roles | Drop CCC User

Centurion User Browser Info as of Mar 20, 2013 12:00

Centurion Logins By Browser				
Browser	Version	Logins	Percent	
MSIE	All Versions	1,337,693	67.83%	
Firefox	All Versions	265,857	13.48%	
Chrome	All Versions	231,463	11.74%	
Safari	All Versions	116,019	5.88%	
AOL Browser	All Versions	14,207	0.72%	
Mobile Safari	All Versions	3,725	0.19%	
RealPlayer	All Versions	2,219	0.11%	
Opera	All Versions	808	0.04%	
Netscape	All Versions	59	0.00%	
NetFront	All Versions	33	0.00%	
BlackBerry	All Versions	27	0.00%	
IceWeasel	All Versions	24	0.00%	
WebTV	All Versions	1	0.00%	
Nintendo 3DS	All Versions	1	0.00%	
ELinks	All Versions	1	0.00%	

Select Survey

Top 10 Centurion Browser Versions				
Rank	Browser	Version	Logins	Percent
1	MSIE	8	609,950	30.88%
2	MSIE	9	548,611	27.82%
3	MSIE	7	155,452	7.88%
4	Chrome	24	78,931	4.00%
5	Firefox	18	72,415	3.67%

### 第3章 調査実施ステージにおける情報通信技術の活用

#### 2. アメリカ労働統計局(United States Bureau of Labor Statistics)

アメリカ労働統計局は、労働省の下で労働に関する統計情報を提供する機関である。ここでは、労働力に関する代表的な統計調査である CES(Current Employment Survey)における取組を整理する。

##### ・CES(Current Employment Survey: 雇用状況調査)における取組

###### (1) 調査概要

調査名	: Current Employment Survey
概要	: 従業員の就業時間、賃金を把握する調査
対象	: 事業所・企業
実施時期	: 毎月実施(毎月 12 日を基準日とする)
回答方法	: EDI(Electronic Data Interchange: 財務情報等のデータをシステムを通じて一定の様式で抽出、送信するもの) web によるオンライン調査 CATI(Computer-Aided-Telephone-Interview: 調査員による電話調査において、PC を用いながら質問、回答記録を行うもの)など
抽出方法	: 層化抽出(民間事業所・企業)
回収数	: 約 14 万 1 千事業所・企業(月あたり)

現在の CES は、1915 年に開始された 4 つの産業(靴、綿製品、綿製品の仕上げ、靴下・下着)における月次の雇用状況調査に由来している。CES によって収集された回答データは、44,000 以上に及ぶデータ系列を毎月作成するために用いられており、雇用、就労時間、賃金、時間外就労時間、女性就業率といった指標が含まれている。このような指標は、景気動向や所得傾向などを把握するために重要な指標であり、他の経済指標(Industrial Index: 工業生産指数、Index of Leading Economic Indicators: 景気先行指数、など)の算出にあたっても参照されるものである。また、CES データは、雇用情勢(The Employment Situation)、実質賃金(Real Earnings)、連邦・州レベルの雇用状況(Regional and State Employment and Unemployment)、大都市における雇用状況(Metropolitan Area Employment and Unemployment)といった毎月公表される統計情報にも利用されるため、毎月のオペレーションを迅速に行う必要性が高いものである。

1984 年には、データ回収の効率性向上に向けて、情報通信技術の活用検討が開始された。現在では、CATI、EDI、TDE 及び web 回答の 4 つが、主要な回答方法となっている。

###### (2) 情報通信技術の位置付け

CES では、回収率を向上させるべく、情報通信技術を活用した複数の回答方法が提供されている。CES は回答が義務付けられていない調査であるため、可能な限り高い回収率を実現する必要性が高い。同時に回答者の回答負荷を下げるために、情報通信技術が活用されている。

一方、BLS にとっては以下のような利点を得ることが可能となっている。

- データ品質の向上
- 必要な時間の短縮
- 回収コストの削減
- 途中脱落の削減
- 他の統計調査との連携可能性

初めて調査に参加する回答者に対しては、事前の依頼状が郵送・FAX で送付される。多くの場合、CATI による回答を 5 カ月続けた後に、自記入式の回答方法(FAX やインターネット回答)へ移行することが調査員の判断に基づいて推奨される。ただし、回答履歴の浅い回答者に対する恒久的な回答方法として CATI が位置付けられている。

(3) 組織体制・機能

調査で利用されているCATIのソフトウェアは、BLS内で開発された。blaiseを利用することも検討されたものの、柔軟性に欠けていたため、インハウスでの開発が決定された。

なお、回答データの収集は各地に設置された下記5か所のセンターが担っている。

図表10 CES のデータ収集センター

ダラス(テキサス州)
フォートウォルトンビーチ(フロリダ州)
アトランタ(ジョージア州)
カンザスシティ(ミズーリ州)
シカゴ(イリノイ州) ※EDIのみ

シカゴ以外のセンターでは、電話による回答の督促も行われている。勤務する調査員は、人材派遣会社を通じて採用されており、現在約300名のCATI調査員が勤務している。調査員は契約社員として雇用されている。

なお、ハードウェア(電話設備など)はBLSの所有物として管理されている。

(4) 情報通信技術の導入経緯・目的

CESが開始された1915年から1983年にかけての主な回答方法は、郵送による調査票の送付と返送であった。ある月の結果が記入された調査票がBLSに返送された後、BLS職員によって調査票の回答結果が入力され、その調査票を対象者に再返送、1年間の回答を記録する形であった。

CESの回答方法、および調査実施ステージにおける情報通信技術は、以下のような変遷をたどってきた。

図表11 CESにおける回答方法の概要一覧

回答方法・導入年	詳細	対象事業所	2011年における回収割合	回答督促の方法	2011年における回収コスト(1サンプルあたり)	2011年における初期回収率
CATI 1984年	毎月、調査員が回答者に電話をかけて質問 回答データは同時に入力されるとともに必要に応じて修正される	初めて回答する企業、自記入を希望しない企業	18%	-調査予定日を記載したハガキを毎月12日頃に郵送 -CATIによる電話	\$ 10.38	90.8%
TDE 1986年	回答者が通話料無料の番号に電話をかけて、自動化された質問に回答 回答者は電話のキーボードを使用する	単一の事業所、比較的単純な雇用状況の企業	4%	-毎月12日にFAXによる事前通知 -締切の2~7日前のFAX -電話 -締切日の朝のFAX	\$ 2.88	84.6%
EDI 1994年	回答者は自企業のシステムからのデータを送付 ファイルは電子的に送信・受信される	通常の回答方法が困難な大規模企業	45%	-事前に決められた日にデータが送付されていない場合の電話	\$ 0.50	59.2%
FAX 1995年	毎月、調査票がFAXされる 回答者は調査票に記入しFAXで返送する	事業所が5箇所以上ある企業	5%	-締切数日前に調査票をFAXで送付 -未回答者に対する確認の電話	\$ 5.86	85.8%
Web 1996年	回答者はBLSのWebサイトと接続し回答を入力する 回答終了前にデータは自動的に修正される	事業所が5箇所以内の企業	25%	-毎月12日に送付される事前通知のeメール -締切の2~7日前のeメール -電話 -締切日の朝のeメール	\$ 2.40	78.5%

### 第3章 調査実施ステージにおける情報通信技術の活用

#### 1915年～1983年:郵送による回答とその課題

1990年代初頭までは、CESの回答方法は郵送に限られていた。調査票が毎月郵送されるとともに、回答者にはできる限り早急に回答を記入して返送するように依頼された。一方で、郵送によるやり取りにおいては、誤回答の修正などに時間がかかることや、修正がさらなる修正を必要とする、といった問題が生じた。特に、発見された誤りが回答者の回答時点で発生したものなのか、回収後のデータ入力時に発生したものなのかについて、確認する必要が生じた。回答者に対するフォローアップについては、確認すべき内容の重要性や複雑さに応じて、電話あるいは調査員による訪問によって行われるなど、コストがかかるものであった。

また、郵送による調査票の回収を行うために、毎月の調査票郵送、および返送に伴う往復分の郵送費や、調査票の回答結果入力やフォローアップに係る人件費など、多くの費用がかかることが課題であった。

#### 1984年:CATIの導入

CESにおいては、結果の一次速報を行う都合上、データ回収期間が10～16日間と極めて限られている。そのため、データ回収に必要な時間を短縮することを目指した。加えて、結果の速報から確報までに行われる修正を少なくすることも目指された。

1984年には、CATIの導入に向けて、400サンプルに対する回収テストが行われた。1987年には、より大規模なサンプルに対するテストが行われたが、最終的には5,500サンプルに拡大され、7年間にわたって11の州におけるテストが実施された。その結果、CATIが有力な回答モードとして認められ、1995年までに毎月約1万サンプルの回答データがCATIによって回収された。

CATIによるデータ回収については、特に以下の2点が長所となることが確認された。1点目は回収に必要な時間が短縮されること、2点目は回収に関連する業務の効率が向上すること、であった。CATIを利用する場合においても、回答者宛てに事前の協力依頼としてハガキが毎月郵送されていたが、後日電話がかかってくることで認識されるため、自記入式の郵送調査に比べて、回収期間内により多くの回答が完了する傾向が見られた。また、CATIを利用することで、回答が行われている間にも、リアルタイムでの回答データ訂正(エディティंग)が可能となったことで、回答データ内のエラーも削減された。

また、CATIによるデータ回収の検討と同じ時期に、サンプリング方法の変更が行われた。元々は割付抽出(Quota-based sample)であったものを、確率抽出(Probability sample)へとサンプリング方法が変更された。この見直しを行うと同時にCATIによるデータ回収を本格的に開始することとなったが、それまで課題として認識されていた、回答依頼および拒否の比率を標準化できない点についても、CATIによるデータ回収の一環として行うことができるようになった。

一方、CATIによるデータ回収の短所として、電話による調査を行う調査員費用がかかることが指摘された。CESにおいては、CATIによるデータ回収に関連する費用が、全体の費用の中で大部分を占めることとなった。

#### 1987年:TDE(Touchtone Data Entry)の導入

アメリカ国内において、タッチトーン式の電話が普及するにつれて、CATIよりもデータ回収費用を少なくさせつつ、かつ、郵送よりも高い回収率を実現し得るTDEについて、研究が始められた。1987年には、メイン州とフロリダ州を対象にして合計200サンプルに対するテストを行った。その結果として、TDEによるデータ回収が効果的であることが確認され、1994年から2003年にかけて30%以上のデータ回収がTDEによって行われることとなった。

調査員が介在する CATI から、回答者の自記入による TDE へと回答方法を変更する際は、約半年のトレーニングを回答者が経ることでスムーズな移行を図った。回答者が CATI による回答や回答時期について習熟したところで、TDE による回答が選択肢として提示されるようになっていた。回答内容が複雑である場合や、TDE による回答を望まない場合には、CATI による回答を続けることも認められた。

TDE が初めて導入された際は、通話料無料の電話回線は州ごとに用意され、回答者からの質問に答えるヘルプデスクが設置された。回答負荷を下げるために、回答データの修正は通話が終了した後に行うこととした。そのため、データの修正が完了するまでに若干の時間がかかることとなり、この点が TDE の短所として指摘された。

#### 1989 年:VR(Voice Recognition)の導入検討

TDE による回答が可能な CES の対象者は、1980 年後半頃に至っても、全体の 75～85%程度にとどまっていた。CATI によるデータ回収もキャパシティが限られる中で、さらなる回収率の向上を目指して VR(Voice Recognition 音声認識)についても検討が行われることとなった。TDE 同様に、VR による回答は、CATI による回答に習熟した後に選択肢の 1 つとして回答者に提示された。

検討の中で、VR は回答者からは好意的に評価され、TDE とほぼ同様の回収率およびエラー発生率であった。しかしながら、VR による回答テストが終了するまでも、タッチトーン式電話の利用が増えつつあり、すでに確立されていた TDE による回答方式以外の、自記入式の回答方式についてはそれ以上のリソースを投入しないことが決定された。

#### 1995 年:EDI(Electronic Data Interchange)の導入

企業活動における PC の利用が一般的になり、EDI 技術も普及しつつあった。BLS では事業所からデータを収集する統計調査を行っているが、複数の事業所を持つ企業からは、回答負荷を減らすことが BLS に求められるようになっていた。

そこで CES では、1995 年 2 月に EDI センターを設置し、特に大規模な企業からの回答データを電子的に収集することとした。EDI センターが設置されて間もなく、他の調査による回答データも EDI センターへ提出することができるようになった。

EDI による回答は、大企業を中心とする回答者の回答負荷を減らすことに加えて、回答に係るコストをも減らすことにつながった。BLS の側には、データ品質の向上、回収コストの減少、さらにより多くの情報が収集可能となる、など様々な長所が存在した。EDI センターでは、複数の調査で回答データのフォーマットを統一することで、データ処理に係るコストの削減も図られた。EDI によるデータ回収を行うには、準備期間が長く必要となること、また、システム構築などの関連する費用が必要となるものの、いったん EDI による回答が開始されて定常的な運用が始まれば、回収コストは極めて低いものとなった。

テストデータが品質面で問題ないことを確認した上で、EDI による回答を希望する企業は、月次ファイルを EDI センターに安全な方法で送信することとなるが、しかしながら、EDI の欠点のひとつとして、データファイルで送信される回答内容に関して確認が必要な場合、システム的な問題か否か、といった点をも確認する必要があることが挙げられた。

### 第3章 調査実施ステージにおける情報通信技術の活用

#### 1996年:One Point TDEの導入

1987年に開始されたTDEでは、各州が個別にTDE専用の番号を設定し、サポートを提供していたが、1996年6月には2つの州に集約される形となった。回答の依頼方法やデータ修正のルールが標準化され、業務の効率化と集約化が図られた。

#### 1995年:FAXの導入

CESでは、中小規模の企業からもデータを回収する必要があるが、そうした規模の企業はEDIによるデータ回収には向かない一方で、TDEや通常のCATIでは回答負荷が高い、という点が課題となっていた。そのような中、FAXによる回答が可能な調査票を、特に複数の事業所を持つ企業を想定して作成することとした。回答依頼も兼ねて、空白の調査票が回答者にFAXで毎月送付され、回答記入後にFAXで返送、CATI調査員によって入力された。回答エラーの発生率は、他の自記入式の回答方法とほぼ同一であった。費用の面から見ると、FAXは郵送やCATIよりは低コストであるものの、TDEよりは高コストであった。

なおFAXによって送付された調査票に対して、画像認識技術を利用した読取りも検討されたが、当時はソフトウェアが未発達であったこと、また、人の目による確認が多く必要とされたことから、利用は見送られた。

#### 1996年:Web Reportingの導入

アメリカの連邦レベルで実施される統計調査において、CESはインターネット回答の実験を行った、初めての統計調査であった。タッチトーン式の電話のみが必要となるTDEとは異なり、インターネットによる回収ができるか否かは、回答者がPCにアクセスできるか、インターネットおよびeメールを利用できるかどうか強く依存している。CESではまず7人のTDE回答者に対する小規模なテストから開始し、50人以上の規模へとテストを拡大、1998年までにCESではインターネットによる回収を行うことを決定した。

TDEと比較した際のインターネット回答の最大の長所として、前月の回答結果が表示可能となる点が挙げられる。これによって、回答している途中から、より複雑なデータ修正が可能となった。加えて、記入がされていない項目についても注意が促されることで、個別の項目に対する回答率も、他の回答方法よりも高くなり、データ修正の必要性が少なくなった。また、回答依頼もeメールで送信されることから、インターネット回答がBLSにとって最も費用効率が高い回答方法となった。

CESがインターネット回答に対する取組の先駆者であったことから様々な試行錯誤が行われた。インターネットによる回答にはシステムにログインするためのパスワードや電子認証が必要であったため、回収率はTDEよりも低かった。このことはより安全な環境を提供することが、インターネット回答を敬遠する結果となる、という課題を残した。

#### 2004年:One Point TDEの導入

2004年7月までに、すべての州がTDEに関連する業務を、BLSが運営する1箇所のTDEシステムに統合することに合意した。このことによって、各州で個別に行われていたヘルプデスク運営や確認のための電話といった業務が集約され、効率的な運用が行われるようになった。

### 第3章 調査実施ステージにおける情報通信技術の活用

#### 2004年:IDCF および Web-Lite(インターネットによるデータ回収設備)の導入

2004年には、CESにおけるインターネット回答は、BLSが管理する共通のプラットフォームであるIDCF(Internet Data Collection Facility:インターネットデータ回収設備)に集約された。プラットフォームを共通化することで、異なる調査においてもセキュリティや業務オペレーションといった共通した部分の管理が容易になった。また、Web画面表示についても見た目や操作感覚が標準化されるとともに、個々の回答者は同一の接続セッション内においてBLSが実施する他の調査に回答することも可能となった。IDCFを利用したBLSによる調査が増えるにつれて、1回答者あたりのコスト削減につながった。開始当時は、同じ自記入式であるTDEよりも回収率が低かった。

2006年には、回収率を高めるためにパスワードを必要としない新しいWeb回答サイトのテストを行った。“Web-Lite”と名付けられた回答サイトでは、回答者は9ケタのCES回答者番号とCAPTCHA認証(Completely Automated Public Turing Test To Tell Computers and Human Apart: データ入力がコンピュータによるものではなく人間によって行われていることを確認する)のみが必要とされた。テストの結果から、Web-Liteを利用するほうが回収率が高く、また、回答者の満足度も高いことが確認された。

しかしながら、回収率とセキュリティはトレードオフの関係にあることも明らかになった。CES回答者番号のみが入力され、先月の回答結果の表示や連絡先情報を表示することができず、当月と先月の回答間での整合性チェックを回答中に行うことができなかった。加えて、回答者はCES側が正しい情報を持っているか否かを確認することができず、連絡先情報を更新することも困難であった。こうした短所もあったものの、CESではセキュリティを多少は犠牲にしてでもより高い回収率を目指すことを優先した。

#### 2005年:インターネット回答の促進

2005年には、CESの月次調査項目に複数の項目(全従業員の勤務時間と賃金など)が追加された。この追加によって、回答負荷がほぼ2倍に増加したため、多くの回答者にとってTDEで回答するには長すぎ、かつ負担が大きくなった。結果として、BLSではTDEによる回答を行っていた回答者を含めて、多くの回答者に対して回答方法をWebへ移行するように促すこととなった。ただし、TDEも回答方法の選択肢の1つとして希望すれば利用できるようにされていた。

#### 2006年: eメールの導入検討

CESによるWeb Liteテストの一方で、eメールによる回答データ収集についても検討が行われた。しかしながら、メールソフトウェアの違いによって、htmlの読み込みが異なることから、回答者によっては画面が正しく表示されない、という問題が生じた。CESでは、eメール経由でWebサイトへアクセスすることで対応を図ったものの、回答者の手作業による部分が多く発生したことから、eメールについてはそれ以上の検討は行わないこととした。

### 第3章 調査実施ステージにおける情報通信技術の活用

#### 2007年:Web ftpの導入

2007年には、回答データ収集およびデータ・エディティングに関する業務をBLSが一括して担うこととなった。ウェストバージニア州においては、エクセルファイルによる回答が中規模企業向けに行われていたが、そのエクセルファイルをテンプレートとして、統一された様式によるWeb ftp方式での回答が始められた。Web ftpによる回答は、事業所が5箇所以上100箇所未満の企業に限定された。回答ファイルはWeb Liteと同様のWebサイトに対してアップロードされる方法であった。

中規模企業の回答者にとっては、エクセルは使い慣れたソフトウェアであったが、そのためにスプレッドシートを修正してしまい、集計時に問題を引き起こす場合もあった。Web ftpは、提出されたスプレッドシートの様式が変更されていない場合には極めて効率的であったが、問題があった場合にはその原因を突き止めて修正するために大きな労力が必要とされた。加えて、回答者の中にはスプレッドシートの中に、コメントを記載し、返答を期待することもあったが、そうしたコメントも確認されない場合があった。

#### 2008年:郵送による回答の減少

サンプリングが確率抽出(Probability sample)となって以降、回答者にとっては郵送に依らない回答方法がより高い回収率につながり、BLSにとってもよりコスト効率が良い回答方法であることが明らかになってきた。CESでは、回答方法としての郵送を推奨することを取りやめるとともに、郵送を最後の手段として位置付けることとした。

2008年には、郵送による回答を行っていた回答者についても他の回答手法へと移行し、現在では郵送は最初の依頼時と移転時の登録、ハガキによる督促に限って用いられている。CESにおける回答方法別のサンプル数割合は以下のような推移となっている。

図表12 CESにおける回答方法別サンプル数割合(%)

	1993年	2003年	2004年	2005年	2006年	2007年	2008年	2009年	2010年	2011年	2012年	2013年
郵送	86	4	3	2	2	2	1	0	0	0	0	0
CATI	4	16	20	22	23	25	16	19	19	17	21	26
TDE	8	31	27	24	23	16	13	8	6	4	3	3
FAX	0	14	14	15	13	8	9	9	7	5	4	5
WEB	0	0	1	1	2	7	13	16	20	27	25	21
EDI	0	29	30	31	33	39	45	44	45	44	43	42
その他	2	6	5	5	4	3	3	4	3	3	3	3

※「その他」について:ごく少数ではあるものの、エクセルファイルによる提出やフロッピーディスクなどを利用してデータを提出する事業所が存在する。

### 第3章 調査実施ステージにおける情報通信技術の活用

なお、CES では推計結果を3回に分けて速報しているが、それぞれの速報における回収率は以下の通りとなっている。

図表13 CES における回収率(%) (2013年)

	1月	2月	3月	4月	5月	6月	7月	8月	9月	10月	11月	12月
初期回収率	72.4	76.5	80.0	75.5	82.1	76.6	74.1	72.4	77.9	83.5	76.1	80.6
二次回収率	93.4	93.7	94.1	94.4	96.4	96.2	93.0	93.5	95.0	96.2	94.5	95.9
最終回収率	95.4	97.0	97.2	97.2	97.2	97.0	97.0	95.9	96.7	97.0	97.0	96.8

### 第3章 調査実施ステージにおける情報通信技術の活用

#### (5) 情報通信技術の活用に関する評価

CES では、回収率を高めるために、現在の方法に至るまでに、回答入力の支援や回答依頼方法の改良を多年にわたって行ってきた。しかしながら、全ての企業が、CATIから始め、調査に慣れてから他の回答方法へと移行するため、回答モードによるバイアスはないと考えられている。時系列での比較可能性にも影響は無いものと判断されている。言語面での対応についても、Web では2011年から、TDE では2012年からスペイン語対応がなされている。

回収コストについては、回答方法別に見たコストは以下の通りとなっており、EDI に次いで Web の回収コストが低く抑えられている。

図表14 CES における回収率と回答方法別コスト(2011年平均)

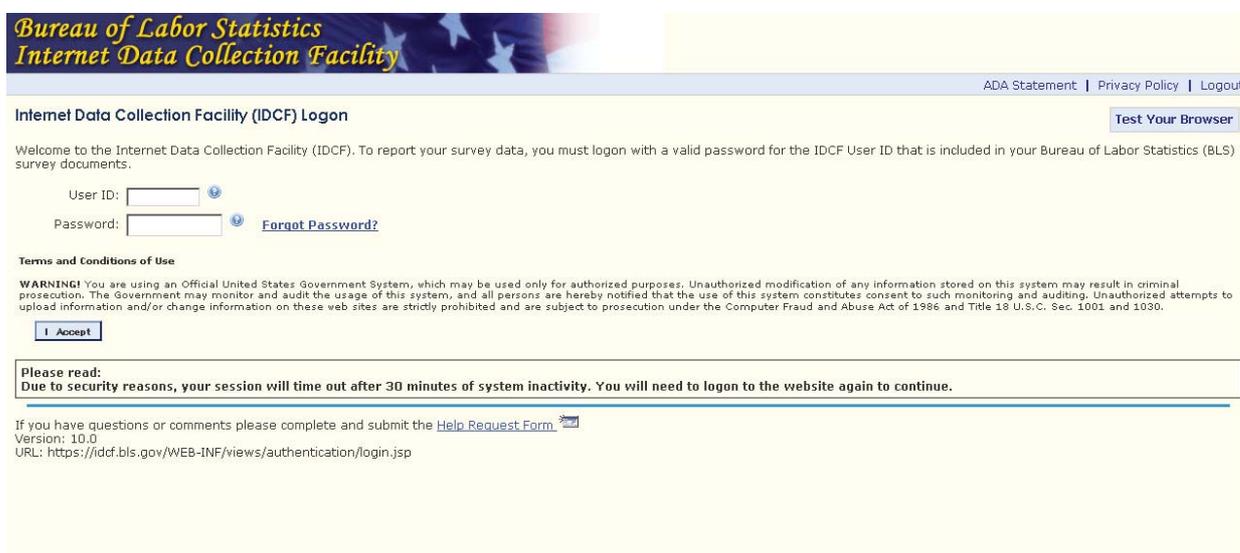
	初期回収率(%)	1回収あたりコスト (USD)
CATI	90.8	10.38
TDE	84.6	2.88
EDI	59.2	0.50
FAX	85.8	5.86
Web	78.5	2.40
その他	-	-

(参考)インターネット回答のインフラストラクチャーIDCF(Internet Data Collection Facility)について  
 BLS では、BLS 全体の取組として、インターネットを利用した回答のためのインフラストラクチャー IDCF(Internet Data Collection Facility)を 2000 年より運用している。以下のようなきっかけをもとに、インターネット回答について本格的な検討が行われることとなった。

図表15 インターネット回答が検討されたきっかけ

インターネットの普及	企業を始めとして一般家庭においても、インターネット接続環境が普及した
インタラクティブ性の向上	Java 技術などの進歩によってユーザーインターフェースがより洗練されたものとなった
セキュリティの向上	セキュリティを確保した環境での通信や情報管理が可能となった

図表16 IDCF のログイン画面



BLS が実施する CES では、IDCF の前身となる web ベースの回答インフラストラクチャーを 1996 年に運用開始した。初期のバージョンでは、比較的小規模の回答者を想定して設計された小規模なものであった。しかしながら、インターネット回答に対する関心が BLS 内で高まるにつれて、BLS 全体として取組む必要性が認識された。特に、データ品質の向上、時間の短縮、コストの削減といった点から、インターネット回答への関心が高まった。その結果、BLS として複数の調査で共有可能なインターネット回答の仕組みを構築することとなった。その決定にあたっては、下記2つの点が目指された。

- 個別の統計調査プログラムに関わらず、BLS を含めたすべての省庁にとっても、様式(見た目や操作感覚など)が統一されたデータ回収のためのインフラストラクチャーとなること

事業所を対象とする CES を含めて、BLS が実施する異なる統計調査に対して、同一の事業所が回答する場合もあることから、様式の統一による回答のしやすさが意識された。

- 共通のインフラストラクチャーを利用することに加えて、個別の統計調査が独自にデータ回収の方法を構築できること

対象者や内容、構成の複雑さなど、統計調査には個別特有の事情が存在することから、データ回収においてそれらの事情を考慮することの重要性が認識されていた。

### 第3章 調査実施ステージにおける情報通信技術の活用

#### [IDCF の特徴]

上記のような思想に基づいて構築された IDCF には、以下のような特徴がある。

- BLS が実施する統計調査におけるポータル(唯一の入り口)として機能する
- 異なる統計調査でも、共通した見た目・操作感覚を持つ
- 複数の統計調査に回答する回答者の支援を行う
- 異なるセキュリティレベル設定を可能とする
- BLS 内部のデータと遮断された状態を維持するファイアーウォールを持つ
- BLS によるアクセス管理を行う
- モニタリングやリスク管理を行うプラットフォームとして機能

インターネット回答のプラットフォームである IDCF は、現在以下の 9 つの調査で利用されている。CES が 1996 年に開始したのに続いて、2003 年には SOII でもインターネット回答が始められた。

図表17 IDCF を利用する統計調査

-Current Employment Survey (CES)	-International Prices (IPP)
-Multiple Worksite Report (MWR)	-Producer Prices (PPI)
-Job Openings and Labor Turnover (JOLTS)	-Occupational Injuries and Illnesses (SOII)
-Annual Refiling Survey (ARS)	-National Compensation Survey (NCS)
-Occupational Employment Statistics (OES)	

前述のとおり、CES において IDCF を利用して回答を行うには、CATI による回答を通じて CES の調査内容に習熟することが求められている。回答者が、IDCF を通じたインターネット回答を望んだ場合、インターネット回答に必要な ID とパスワードに加えて、下記のような説明書(インターネット回答についてのパンフレット、IDCF 上でのアカウント作成・ログインについてのパンフレット)が郵送される。

図表18 CES インターネット回答のパフレット

U.S. BUREAU OF LABOR STATISTICS

Current Employment Statistics  
Reporting CES Data  
on the Internet

u have any q

**1-800-827-2005**

---

**Reporting CES data on the BLS Internet**

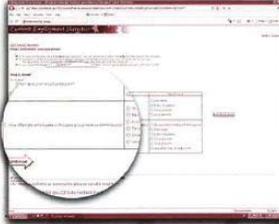
U.S. BUREAU OF LABOR STATISTICS

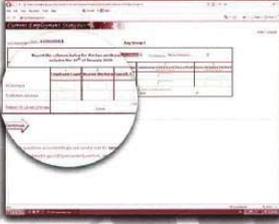
- 1** Link to <https://cesdata.bls.gov>  
Please note that the "s" in the https is required.  
If you experience any difficulties, send an email to [CESWebHelp@bls.gov](mailto:CESWebHelp@bls.gov) or call 1-800-827-2005. This mail link is at the bottom of every page. You can click on this link if you need additional assistance.  
Enter your **CES 9-digit report number**. If you can see the letters in the box, enter the letters and click "I accept." If you can't see the letters, click on "Can't see the letters?" It brings up another screen.


- 2** The next screen contains a welcome message.  
**Click Continue.**


- 3** Enter your contact information.**Click Continue.**  
This screen will only be displayed the first time you log on to the system.  
To update your information in the future, **Click Update Respondent Information** at the upper-right hand corner of the welcome screen.


- 4** Select the pay groups and frequency  
**Click Continue.**


- 5** Enter your information!  
**Click Continue.**


- 6** Review the data you entered.  
**Click Submit Data to BLS.**

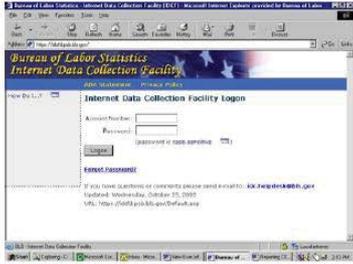


図表19 IDCF でのアカウント作成のためのパンフレット  
(ログイン方法とアカウント作成の方法が記載されている)

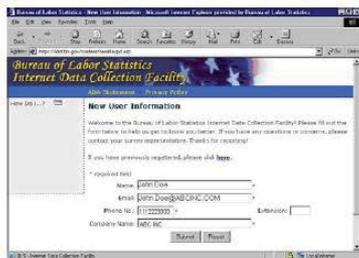
**Establishing a BLS Internet Account for CES/MWR File Transfer**  
Before you can report your Bureau of Labor Statistics (BLS) data through the Internet, you must first establish a BLS Internet account. To do so, please follow the steps below.

1. **Website Address:** Go to <https://idcf.bls.gov>
2. **Account Number/Password:** At the *Internet Data Facility Logon* screen enter the 12-digit Temporary Account Number and Temporary Password from your cover letter or form and click **Logon**.

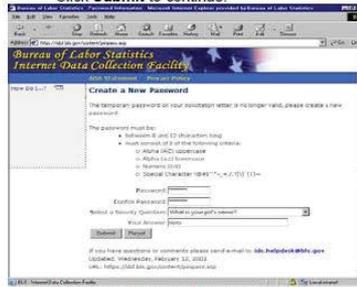
If you do not have a Temporary Account Number and Password, please contact the help desk at [mwr\\_helpdesk@bls.gov](mailto:mwr_helpdesk@bls.gov) to receive one.



3. **Contact Information:** At the *New User Information* screen, enter all of the required information (items denoted with an asterisk, [\*]) and then click **Submit** to continue to the next step.



4. **Password:** At the *Create a New Password* screen, enter a new password that you will use to access your account. Do not enter the temporary password you were provided. **Your password must be 8-12 characters in length and must be a combination of upper and lower case letters and numbers. Examples: Bisdats9, Boston01.**
5. **Security Question:** In the "Select a Security Question" box, please select one of the four questions. Enter your answer in the "Your Answer" box (maximum of 30 characters). **Note: In case you forget your password, the system will allow you to retrieve it if you answer the Security Question correctly. Click Submit to continue.**



6. **Permanent Account Number:** Your permanent BLS account number will be displayed on the *Account Information* screen. **Please make a note of the account number or print this page for your reference.** You will need the account number when reporting your data through the BLS Internet Data Collection Facility.



7. **Account Established:** Your BLS Internet account has now been established. If you are ready to transfer your file, click **Enter Data** and follow the instructions in the *Reporting CES/MWR Data using Internet File Transfer*. If you are not ready to transfer the file, please exit and return to <https://idcf.bls.gov> when you are ready to do so.

### Reporting CES/MWR Data Using Internet File Transfer

1. Create your CES/MWR data file as specified by BLS. For ease of identification, include the reference month and year in the file name. For example: **ABCINC0403.BLS**
2. At the *Internet Data Collection Facility Logon* screen, enter your permanent account number and password and click **Logon**.





## 第3章 調査実施ステージにおける情報通信技術の活用

### 3. カナダ統計局(Statistics Canada)

カナダ統計局は、同国における統計調査の企画・実施を行う機関である。

ここでは、代表的な労働力に関する統計調査である LFS(Labour Force Survey)における情報通信技術の活用に関する取組を整理する。特に LFS では、EQ(Electronic Questionnaire: 電子調査票)と呼ばれるインターネットによる回答についてパイロットテストを実施しており、そこでの検討内容について整理を行う。

#### ・LFS(Labour Force Survey: 労働力調査)における取組

##### (1) 調査概要

調査名	: Labour Force Survey
概要・目的	: 個人の就労状況を把握する調査
対象	: 15 歳以上の個人を対象
実施時期	: 毎月 15 日を含む週を基準週として翌週に実査を行う
調査手法	: CATI 及び調査員による訪問面接(ノート PC を用いた CAPI) 電話番号が不明の場合には、初回は訪問した上で、残る 5 か月間については CATI を用いる。
抽出方法	: 全国(10 州に加えて 3 準州を含む) 層化抽出された世帯対象者が 6 か月間にわたって回答を行うパネル調査 (毎月新しいパネルと入れ替えるため、1/6 の回答者は新しい対象者となる)
回収数	: 約 5 万 4 千世帯 約 10 万個人

LFS は、雇用・失業に関する統計調査である。データ収集後、13 日後には結果が公表されており、月次の統計として迅速な情報提供が行われている。LFS の主な目的は、生産年齢にあたる人口を「雇用状態」、「失業状態」及び「非労働力人口」の 3 つに区分することで、労働力市場の状態を把握することにある。一般的に知られている失業率を算出する基礎となるほか、就業者人口や雇用率といった指標を提供する調査である。加えて、産業別、職業別、公共部門、民間部門といった区分での就業者人口、労働時間等も提供している。主な質問項目は、雇用・失業状況、労働時間、産業、職業、賃金等である。

LFS は 1945 年に四半期ベースの調査から開始され、1952 年には月次ベースでの調査となった。

##### (2) 情報通信技術の位置付け

LFS では、回答方法の選択肢を増やすことで、回答者にとっての利便性が高まり、回答率の向上につながると考えている。また、調査員の旅費を含め、データ回収コストを削減するために、最初のコンタクトでは CAPI を用いるものの、その後は CATI へと移行することを推奨している。毎月の回答は 10 日間で、最初のデータ回収を”birth”と呼び、電話あるいは調査員による訪問が行われる。次の回答では、多くの場合電話調査が行われる。

後述する EQ(Electronic Questionnaire: 電子調査票)は回答者にとってより便利で、かつ調査員数も少なくて済むため、効率的な業務につながると考えられている。

##### (3) 組織体制・機能

カナダ統計局では、以前は 5,000 名程度の職員数で、現在は 4,300 名ほどと削減されてはいるものの、システムの開発を行うプログラマーや統計学に精通したスタッフ、アナリスト等が所属しているため、パイロットプロジェクトもカナダ統計局内で完結させることが可能であった。

#### (4) 情報通信技術の導入経緯・目的

1945年に開始されたLFSの回答方法は、当初は紙の調査票からCAPI、CATIへと変遷を遂げてきた。1980年代にはPCと電話回線によるデータ送信技術を利用してCAPI、2000年頃にはCATIを導入した。CATIでは、ソフトウェアとしてblaiseを利用している。

2004年までは、最初のデータ回収は調査員が訪問するCAPIのみを利用していたが、現在ではCATIも利用されている。

自記入式の回答方法を導入した最大の理由は、コスト削減であった。調査員の人件費や教育費用を削減することが可能となったが、特にLFSは毎月実施される調査であるため、コスト面でのメリットが大きかった。一方、カナダ政府の規制基準に適合させる必要があったため、ソフトウェアはblaiseからインハウスで開発したソフトウェアに移行することとした。機能面で優れた商品は多々存在していたものの、いずれも基準に添っていないため利用することができなかった。

ここからはEQによるパイロットテストについて整理を行う。

##### ①パイロットテストの概要

ターゲットは6,400世帯で、半数(3,608ケース)が2013年3月に、残り半数(3,601ケース)が2013年4月に開始された。テスト期間は6月までの3カ月あるいは4カ月であった。対象は4つの州(ニューブランズウィック州、ケベック州、オンタリオ州、アルバータ州)とし、オンタリオ州では都市部を中心とする構成とした。

##### ②EQの対象条件

EQに回答することができる回答者には、以下のような条件が設定された。いずれも回収データの正確性を期すためのものである。

- 同居世帯の中で就労者が1名のみ
- 世帯が6ケタの郵便番号を持ち、かつ有効な電話番号を持つ
- 現在の世帯構成員がパイロットテスト期間中継続してコンタクト可能
- 世帯構成員による回答を通じて回答完了が可能
- 実在しない世帯構成員が含まれない

##### ③EQによる回収

最初のデータ回収では通常と同じく、電話あるいは調査員による訪問が行われるが、質問の最後に、翌月以降の回答をインターネットで行いたいかなどが確認された。翌月以降、EQを利用する場合、回答期間は3~4日間となり、回答がない場合は電話による回答依頼を行った。

### 第3章 調査実施ステージにおける情報通信技術の活用

#### ④パイロットテストによる検証内容

EQ によるパイロットテストでは、下記の内容について検証が行われた。

図表21 パイロットテストによる検証内容

<p>[オペレーション上の実現性評価]</p> <ul style="list-style-type: none"> <li>-EQ の回答期間の長さ</li> <li>-e メールによる回答依頼と督促による回収率への影響</li> <li>-回答パターンに対する影響(電話や調査員による訪問調査とEQ の比較)</li> <li>-データ回収コスト(ヘルプデスクやコーディング、フォローアップなどを含む)</li> <li>-部分非回答や代理回答への影響</li> <li>-回答者からのフィードバック</li> </ul>
<p>[ソフトウェアの開発]</p> <ul style="list-style-type: none"> <li>-集計システム(データ・エディティング、補完推計など)</li> <li>-EQ への移行依頼</li> <li>-調査票の設計</li> </ul>
<p>[データ品質に対する影響評価]</p> <ul style="list-style-type: none"> <li>-非回答など(面倒な質問を飛ばすなど)</li> <li>-回答モードによるバイアス、推計への影響</li> </ul>

データ品質に対する影響評価では、以下の項目を中心に検証が行われた。

図表22 データ品質に対する影響評価において検証が行われた項目

EQ 回答者の属性	非回答(スキップ)
推計	就労状態の変化
産業および職業の分布	回答の分散
定性的なフィードバック	回答の平均値
回収率	失業率の推計
世帯構成員数の変化	
補完推計および代入	

パイロットテストと実際の調査結果を比較したところ、推計には影響を及ぼさないことが確認された。具体的には4つの地域において17項目の推計を比較、サブグループとして5つの性年代のグループについて検証を行ったが、いずれも有意差は見られなかった。

なお、回答者から挙げられた定性的なフィードバックは以下の4つのカテゴリが多かった。

- ログインの方法やパスワードに関連するもの
- 回答期間が短い
- 個別の質問に関するもの
- EQ の依頼に関するもの(カナダの法律によって依頼文が長文になっていた)

#### (5) 情報通信技術の活用に関する評価

EQ のパイロットテストからは以下のような示唆が得られた。

図表23 パイロットテストからの示唆

- 回答期間は長いほうが好ましい
- EQ による回答に同意した回答者は、他の回答方法でも回答する協力的な回答者であることが多い  
(CATI や CAPI に抵抗感を持つ回答者にとって好ましいと想定していたが、実際は異なっていた)
- リマインダーメールの効果が高い  
(回答の送信時間帯を分析すると、リマインダーが送られた月曜と水曜にもピークが存在していた。毎週土曜の夜に送信される最初の e メールより、リマインダーのほうが効果的であることが示唆された。)
- CATI 調査員の作業負荷が削減された  
(1 か月分の回収あたり 92 時間分の作業時間の軽減につながった。)

#### (6) 課題と今後の方向性

今回パイロットテストが行われた EQ のソフトウェアは、CAPI/CATI のソフトウェアと異なるものであった。そのため、EQ のパイロットテストでは、回収したデータを blaise のフォーマットに修正していた。

カナダ統計局では一つのサーバーで、全ての回答方法をサポートし、データ管理を一括してできるアプリケーションとして、Integrated Collection and Operation System (ICOS) を開発している。開発にあたってはカナダ統計局内で行う予定であり、3 つの異なる回答方法を 1 つのアプリケーションで管理することで、さらなるコスト削減を目指している。なお、blaise も CAPI/CATI とウェブ回答を統合したシステムを公開すると見られているが、カナダ統計局では blaise は利用しないことをすでに決断している。

ICOS の導入によって、CAPI や CATI で回答している間は、ウェブ回答を不可能とするなど、実査管理の面からも効率化が期待されている。同じく、ウェブ回答中は CAPI/CATI のアクセスが不可能とする。また、データ送信も従来よりも簡単になる見通しである。

カナダ統計局では 2015 年を目標として、すべての調査に対して EQ を導入することを検討しているものの、2015 年以降にずれ込む可能性が高い。一方で、EQ では、協力的な回答者の回答が回収されるため、調査員にとっては回収が困難な回答者が残される、といった課題も指摘されている。

なお、システム開発に費やしたコストは、EQ の活用が進むにつれて十分に回収ができると考えられており、さらなる EQ の活用が目指されている。

### 第3章 調査実施ステージにおける情報通信技術の活用

### **第4章 結果活用ステージにおける情報通信技術の活用**

本調査研究では、結果活用ステージにおける情報通信技術の活用に関する取組事例として、国内及び海外において、どのような取組がなされているかの整理を行った。

なお、海外については、「統計データの二次利用促進に関する研究会 報告書」(平成20年10月統計データの二次利用促進に関する研究会)においても、統計データの結果活用の実態や今後の課題について整理・検討が行われている。そのため、本報告書では情報通信技術に着目した整理を行うこととした。

具体的には、カナダ統計局、シカゴ大学 NORC、ミシガン大学 ICPSR の3つの機関における取組の整理を行うこととした。

## I 国内における情報通信技術の活用に関する取組

### 1. 統計センター(e-Stat)

#### 1) e-Stat の概要

統計センターでは、各府省情報化統括責任者(CIO)連絡会議で決定された「統計調査等業務の業務・システム最適化計画」に基づき、日本の政府統計関係情報のワンストップサービスを実現するために、政府統計のポータルサイト「政府統計の総合窓口(e-Stat)」を平成20年度から本運用している。e-Statは、従来、各府省等のホームページに掲載されていた各種統計関係情報をワンストップで提供することを目指しており、各府省等が登録した統計データ、公表予定、新着情報、調査票項目情報等の各種統計情報を、インターネットを通して利用することが可能である。

ここでは、「政府統計の総合窓口」(以下、e-Stat)における取組を整理する。e-Statの取組は、主に①API機能による統計データの高度利用環境の構築、②統計GIS機能の強化、③オンデマンドによる統計作成機能・方策の研究、の3点である。

#### 2) 取組の概要

##### ①API機能による統計データの高度利用環境の構築

e-Statでは、提供している統計データを機械判読可能な形式(XML等)で取得できるAPI(Application Programming Interface)が提供されている。具体的には、以下の5つの機能である。

##### 統計表情報取得(GET)

統計表の情報(統計表ID、調査名、統計表名、調査年月等)を提供する機能。検索キーワード等を指定することで、絞り込みが可能。

##### メタ情報取得(GET)

統計表(統計表ID)に含まれるメタ情報(集計事項、地域事項、分類事項等)を提供する機能。

##### 統計データ取得(GET)

統計表(統計表ID)に収録されている統計データ(数値データ)を提供する機能。必要に応じて、データセット、メタ情報による絞り込みを行うことができる。提供するデータが大量の場合は、分割して提供される。

##### データセット登録(POST)

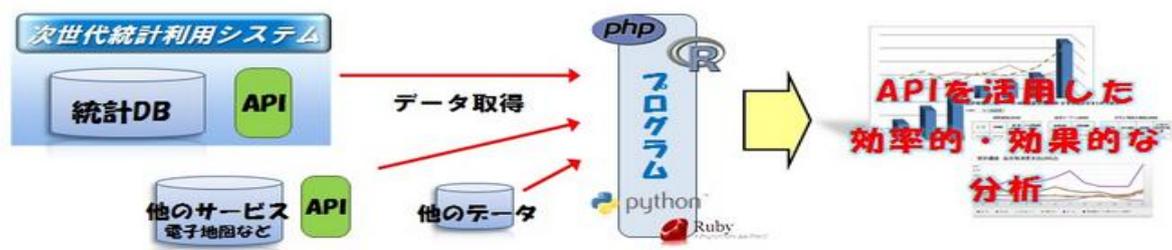
統計データ取得時の絞り込み条件を登録する機能。統計データ取得時にデータセットの取得条件をさらに絞り込むことも可能。登録したデータセットは他のユーザーに公開することも可能。

##### データセット参照(GET)

登録されたデータセットの情報を参照する機能。

APIを利用したアプリケーションを開発することにより、統計データを利用した高度な利用環境を構築することを目指している(試行運用中)。

図表1 API機能のイメージ



(出所: 統計センター次世代統計利用システム「<http://statdb.nstac.go.jp/system-info/api/>」)

### ②統計 GIS 機能の強化

e-Stat では、デジタル化された地図データと、統計データや位置、空間に関する情報などを統合的に扱う地理情報システム(GIS)の仕組みを活用し、利用者個々のニーズに合わせて各種統計調査の統計データを背景地図と共に視覚化して提供している。本システムを利用して統計地図を作製することにより、統計表を眺めただけでは困難な地理的な分析を視覚的に行うことができ、防災活動等への活用が可能となる。統計 GIS 機能の概要及び収録データは以下の通りである。

#### 基本的操作・機能

- 地図操作—Google Maps と同様の地図操作が可能
- ポイント移動—住所、郵便番号、経緯度
- 計測—距離、経路距離、面積
- サブ地図表示—地図画面の2分割表示
- 案内図—メイン地図の広域地図表示・移動
- 最寄駅検索—半径 5,000m 以内の駅名の表示・移動

#### プロット

- 登録—地図クリック(個別)、ジオコーディング、インポート(経緯度)
- 編集—項目の追加、削除、変更
- 機能—名称検索・表示及びポイント移動

#### エリア

- 登録—地図クリック(多角形、フリー円、同心円、バッファエリア、到達圏(徒歩、車))、インポート(shape)
- 編集—項目の追加、削除、変更
- 機能—名称検索・表示及びポイント移動

#### グラフ

- 統計データ又はユーザーデータによるグラフ作成(小地域、メッシュ)
- プロット、エリアの集計及びグラフ作成
- グラフの種類—界面、棒、円、界面クロス
- 編集—ランク分けの種類、階級及び表示色の設定、変更

#### レポート(Excel 形式)

- シンプルレポート—作成したグラフについてのレポート
- リッチレポート—同心円内のエリアについて、年齢構成等の基本的な分析結果

#### 収録データ

- 国勢調査小地域集計(平成 12 年、17 年、22 年)
- 国勢調査地域メッシュ統計(1Km,500m)(平成 12 年、17 年、22 年)
- 事業所・企業統計小地域集計(平成 13 年)
- 事業所・企業統計地域メッシュ統計(1Km,500m)(平成 13 年、18 年)
- 経済センサス—基礎調査地域メッシュ統計(1Km,500m)(平成 21 年、平成 24 年)

## 第4章 結果活用ステージにおける情報通信技術の活用

図表2 GIS 機能のイメージ



(出所:統計センター次世代統計利用システム「<http://statdb.nstac.go.jp/system-info/gis/>」)

### ③委託による統計の作成等及び匿名データの作成・提供

統計法が60年ぶりに抜本改正(平成19年法律第53号)され、平成21年4月に全面施行されたことで、統計法の基本理念が、これまでの「行政のための統計」から「社会の情報基盤としての統計」へと大きく転換した。この転換を受け、統計センターでは、学術研究や高等教育の分野で公的統計の活用を図るための枠組みである二次的利用制度を、「i)委託による統計の作成等(以下、オーダーメイド集計)」と「ii)匿名データの作成・提供」の二つのサービスを柱として発足させた。しかし、データ分析の切り口によっては個人が特定される恐れがあるなど、運営・制度面及び秘匿処理における課題もあり、慎重な検討を行っている段階である。

それぞれのサービスでは以下のような内容が提供されている。

#### i)オーダーメイド集計

統計センターでは、既存の統計調査の集計項目の分類一覧から項目を選択し、それらを組み合わせた統計表の作成を委託することができる。ただし、オーダーメイド集計は、学術研究の発展又は高等教育の発展に資すると認められる場合であり、かつ定められた利用条件を満たす場合に利用可能となる。

オーダーメイド集計の機能には、各府省が統計センターに委託し統計センターから提供されているもの、各府省等が直接提供しているものがある。現在利用可能な統計データ(図表3)及び利用条件は以下に示す通りである。

## 第4章 結果活用ステージにおける情報通信技術の活用

・利用可能な統計データ(2013年12月時点)

図表3 現在オーダーメイド集計が利用可能な統計データ

所管府省	統計センターによる提供	所管府省による提供
内閣府	<ul style="list-style-type: none"> <li>・消費動向調査</li> <li>・企業行動に関するアンケート調査</li> </ul>	
総務省	<ul style="list-style-type: none"> <li>・国勢調査</li> <li>・家計調査</li> <li>・全国消費実態調査</li> <li>・就業構造基本調査</li> <li>・住宅・土地統計調査</li> <li>・家計消費状況調査</li> <li>・労働力調査</li> <li>・社会生活基本調査</li> </ul>	
財務省		<ul style="list-style-type: none"> <li>・年次別法人企業統計調査</li> <li>・法人企業景気予測調査</li> </ul>
文部科学省	<ul style="list-style-type: none"> <li>・学校基本調査</li> </ul>	
厚生労働省	<ul style="list-style-type: none"> <li>・賃金構造基本統計調査</li> </ul>	<ul style="list-style-type: none"> <li>・人口動態調査</li> <li>・医療施設(静態)調査</li> <li>・患者調査</li> <li>・毎月勤労統計調査(特別調査)</li> </ul>
農林水産省		<ul style="list-style-type: none"> <li>・農林業センサス</li> <li>・漁業センサス</li> <li>・海面漁業生産統計調査</li> </ul>
経済産業省		<ul style="list-style-type: none"> <li>・経済産業省企業活動基本調査</li> </ul>
国土交通省	<ul style="list-style-type: none"> <li>・建築着工統計調査</li> </ul>	

・利用条件

- 学術研究又は高等教育(学校教育法(昭和22年法律第26号)第1条に規定する大学又は高等専門学校における教育)の用に供することを直接の目的とするものであること
- 学術研究目的、高等教育目的等の用に供することを直接の目的とすること

## 第4章 結果活用ステージにおける情報通信技術の活用

### ii) 匿名データの作成・提供

匿名データの作成・提供として、利用要件を満たした申出者に対して、特定の個人又は団体等が識別できないように加工して作成した調査票情報の利用を一定期間認めている。研究者等は匿名データを利用することで、行政機関等が作成していない統計表の作成のみならず、多変量解析など、マイクロデータに基づく実証分析を行うことが可能となり、また大学・大学院等の教員は、匿名データを活用した講義又は演習等の実施が可能となる。現在利用可能な統計データ(図表4)及び利用要件は以下に示す通りである。

- ・利用可能な統計データ(2013年12月時点)

図表4 現在匿名データの提供が利用可能な統計データ

所管府省	統計センターによる提供	所管府省による提供
総務省	<ul style="list-style-type: none"> <li>・国勢調査</li> <li>・労働力調査</li> <li>・住宅・土地統計調査</li> <li>・全国消費実態調査</li> <li>・就業構造基本調査</li> <li>・社会生活基本調査</li> </ul>	
厚生労働省		<ul style="list-style-type: none"> <li>・国民生活基礎調査</li> </ul>

### ・利用条件

- 統計の作成または統計的研究にのみ利用されること
- 学術研究目的、高等教育目的等の用に供することを直接の目的とすること
- 学術研究の成果、高等教育の内容等が公表され、社会に還元されること
- 匿名データが適切に管理されていること

## II 海外における情報通信技術の活用に関する取組

### 1. カナダ統計局(Statistics Canada)

※「統計データの二次利用促進に関する研究会 報告書」(平成 20 年 10 月統計データの二次利用促進に関する研究会) 資料6 参照

#### 1) 取組の概要

カナダ統計局では、マイクロデータを含め、集計表等を公表しているが、統計データの提供にあたっては、大きく3つの分類に基づいた提供が行われている。

- ①集計表や分析レポートといった統計結果
- ②個別のリクエストによる集計
- ③マイクロデータを用いたデータ提供

③のマイクロデータを用いたデータ提供では、元データから氏名や住所、電話番号等を削除しただけの秘匿性の強いConfidential Microdata Files や、実年齢をコード化(5区分など)するなどの匿名化が図られたPUMFs(Public Use Microdata Files)が提供されている。

前者のファイルには、原則としてRDC(Research Data Centre)からのみアクセスが認められている。一方、PUMFsには、DLI(Data Liberation Initiative:高等教育機関を中心とするデータ利用の普及取組)を含めて、より多くの利用者からのアクセスが可能となっている。また、近年ではCDER(Canadian Center for Data Development and Economic Research)として、秘匿性が高い経済統計の一部についてのデータ分析を行う機能が提供されている。

#### 2) 情報通信技術の導入経緯

データへのアクセスプログラムに対する需要は、学術機関と公的機関の両者によるものであった。DLIとRDCは学術機関から、CDERは政府機関やカナダ銀行といった公的機関からの要求に応える形として設置された。

情報通信技術の活用にあたって、カナダ統計局では、ユーザーの利便性とセキュリティ確保のバランスを意識した取組が行われている。例えば、PUMFsではユーザーの利便性よりセキュリティを重視した内容とする一方で、マスターファイルであるConfidential Microdataでは、ごく限られたアクセス権限の下で、より秘匿性の高いデータを扱うこととなる。

データ利用者の特性(学術研究者、政策研究者など)と利用目的に沿った形で、より多くの人にマイクロデータが活用されるよう、異なる種類のファイルが提供されている。

## 第4章 結果活用ステージにおける情報通信技術の活用

### 2. シカゴ大学 NORC

NORC は、シカゴ大学内に設置されている社会調査等の実施や研究を行う独立研究機関である。NORC ではアメリカ全土をカバーする調査員ネットワークを保有しており、政府機関や教育機関、NPO、企業に対してデータや分析サービスを提供している。調査員は 800～1,000 人程度の規模となっている。各地域のマネージャーによる管理の下、6～12 種類の調査を実施している。

NORC が関わる分野としては、保健、教育、経済、犯罪、エネルギー、安全保障、環境といったものが挙げられる。

#### 1) 取組の概要

NORC では、data enclave と呼ばれるデータ提供のためのプラットフォームを 2007 年 7 月から提供している。NIST (National Institute of Standards and Technology:アメリカ国立標準技術研究所)による支援の下、研究者に対して秘匿性の高いマイクロデータを、より便利に利用できる環境を提供することが目的とされている。data enclave のサーバーは NORC の本部であるシカゴに設置されている。

運営費用は、NIST (National Institute of Standards and Technology:アメリカ国立標準技術研究所)の技術イノベーションプログラム(Technology Innovation Program)や、農務省、学術研究機関であるコッフマン基金などからの拠出を受けている。

セキュリティ上の問題を防ぐために、data enclave ではシンクライアント端末から、データへアクセスすることが可能となっている。data enclave はインターネットから切り離された環境にあり、事前の確認と承認がない限り、データを移動(ダウンロード等)することはできない。事前の確認は、開示チェック(Disclosure Review)と呼ばれ、問題がないことが確認されれば、データがダウンロード用のサイトに移動され、研究者がダウンロードできるようになる。現在、開示チェックにかかる時間は 24～48 時間程度となっている。

#### 2) 情報通信技術の導入経緯

data enclave は、データ作成者および利用者の双方からのニーズに対応する形で開発が進められてきた。データ作成者にとっては、より高い品質のデータを収集するために多くのリソースを投入する一方で、収集したデータ品質を維持し、回答者の個人情報の保護を図りながら、安全な環境の下でデータをより広く利用してもらいたい、という要望が持たれていた。

データ利用者にとっては、より高い品質の分析を行いたい、という要望が持たれていた。

これら双方の要望に対応するために、よりよいデータ、よりよい分析を行うためのプラットフォームとしての data enclave が作られることとなった。データ利用をより便利なものにするべく、リモートアクセスが可能な環境が用意されており、現在、400 人程度の研究者によって利用されている。

data enclave では、リモートアクセスを実現するために、Citrix と呼ばれる商用のサーバー仮想化技術を利用している。また、近年ではビッグデータに対する関心が高まる中、2012 年に大規模データを処理するための設備を導入している。NORC 外部の技術者とも協働しながら、ソーシャルメディア分析など、今後の活用が想定される大規模データへの対応を図っている。

また、tableau と呼ばれる市販のソフトウェアを用いた集計ツールも提供しており、既存技術を活用しながら効率的な取組を行っている。

### 3. ミシガン大学 ICPSR

ミシガン大学 ICPSR は、社会調査の実施、分析等を行う研究機関であり、世界最大規模のデータアーカイブを運営している。ここでは、主にどのような情報通信技術を活用しているかについて整理を行う。

※「統計データの二次利用促進に関する研究会 報告書」(平成 20 年 10 月統計データの二次利用促進に関する研究会) 資料6 参照

#### 1) 取組の概要

ICPSR では、有意義な分析を行うことができる情報をデータに残しながらも、秘匿管理を厳密に行うことが意識されている。そのため、統計データの活用に関連して、秘匿管理を含め、以下の4つの分野ごとにツール(サービス、ソフトウェアなど)が提供されている。

- ①秘匿管理
- ②利用制限データ
- ③データ集計
- ④データ配付

それぞれの分野では、以下のようなものが提供されており、研究者がデータを利用するにあたって必要となる機能をサポートし、利活用可能な環境を作り出している。

図表3 ICPSR が提供・紹介するツール

	概要	費用
<b>①秘匿管理</b>		
Disclosure Review	-開示にあたって問題がないかをチェック -必要に応じて修正方法を示すサービス	有料
QualAnon	-定性情報(コメントなど)を匿名化する(氏名の変換など) -web 上で動作するソフトウェア	無料
<b>②利用制限データ</b>		
Physical Data Enclave	-極めて秘匿性の高い個人情報を含むデータを保管、利用するための施設 -ICPSR の地下室に所在	無料
Security Inspections	-研究施設等のセキュリティについて監査を行うサービス	有料
Security Plan Reviews	-利用制限データを扱うにあたって必要な安全確保計画を監査するサービス	有料
Virtual Data Enclave	-利用者のデスクトップ上で仮想的に利用制限データを扱えるようにする仕組み -データダウンロード等は不可能 -結果の共有機能などがオプションで提供される	有料
<b>③データ集計</b>		
Colectica for Microsoft Excel	-DDI(Data Documentation Initiative)フォーマットに沿って、情報管理を行うソフトウェア	無料
DataUp	-オープンソースによるエクセル・アドインソフトウェア -データの整理等を行う	無料
Google Refine	-クリーニング前のデータに対してデータ変換などのクリーニングを行うソフトウェア	無料
Nesstar Publisher	-データ、メタデータを整備するソフトウェア	無料
<b>④データ配付</b>		
Dataverse Network	-オープンソースによるデータ共有、分析等を行うソフトウェア	無料

## 第4章 結果活用ステージにおける情報通信技術の活用

### 2) 情報通信技術の導入経緯

リモートアクセスを行うための技術として、ICPSR では VMware と呼ばれる仮想化ソフトウェアが利用されている。Citrix や Microsoft 社の Remote Desktop Technology といった類似のソフトウェアが存在する中で、VMware を採用した理由は、ユーザー間のパフォーマンスに影響が少ないことであった。

Citrix ではシステムリソースをユーザー間で共有するため、複数のユーザーが同時に使用すると PC のパフォーマンスが低くなる。逆に、使用者が少ないとパフォーマンスは高まる。VMware では、ユーザーごとにシステムリソースを割り当てる方式を採用しているため、ユーザー間でパフォーマンスに影響を与えることはない。(ただし、誰も使用者がいない時間帯にアクセスしても、各 PC に配分されたメモリー以上の作業は不可能、という短所がある。)

このように、技術の導入にあたって検討すべき点として、ユーザーに対してどの程度のシステムリソースが必要かを判断する点が挙げられている。物理的に存在する PC とは異なり、常にどこからでもアクセスが可能になる、という仮想化技術によって、アクセスの集中によるシステム負荷につながっている。今後の課題の1つとして、ICPSR では、マイクロデータ以外での提供方法についても検討を行っている。

## 第5章 統計の作成、提供等において情報通信技術を活用する際の示唆・課題

前章までに整理を行った、国内外における情報通信技術の活用に関する取組の整理及び有識者研究会におけるディスカッションを通じて、今後の統計調査に対して情報通信技術を活用する際の示唆課題について、以下のように「1. 調査実施ステージにおける示唆・課題」と「2. 結果活用ステージにおける示唆・課題」に分けた整理を行った。

～統計の作成、提供等において情報通信技術を活用する際の示唆・課題 概要一覧～

1. 調査実施ステージにおける示唆・課題	
1)情報通信技術の位置付け	<ul style="list-style-type: none"> <li>－他の回答方法と組み合わせた位置付け－</li> <li>✓情報通信技術を活用した回答方法は、他の回答方法を含めた1つとして位置付けられる。</li> <li>✓これまでの回答方法(紙調査票による回答など)を存続しながら、情報通信技術を活用した新しい回答方法を導入することが必要となる。</li> </ul>
2)情報通信技術の活用目的	<ul style="list-style-type: none"> <li>－回答者にとっての「回答負荷軽減」と統計作成者にとっての「データ品質の向上」－</li> <li>✓情報通信技術の活用によって、調査プロセス全体を俯瞰した取組みとして、コスト削減を意識しながらも、回答負荷軽減及びデータ品質の向上を目指すことが可能となる。</li> <li>✓一方で、回答方法が異なることによるデータ品質への影響の検証も必要とされる。</li> </ul>
3)統計調査の回答者、回答内容を踏まえた活用	<ul style="list-style-type: none"> <li>－同一客体を対象とした定型の内容を対象とした活用－</li> <li>✓回答負荷軽減、データ品質の向上といったメリットは、同一客体に対して定型の内容を繰り返し質問するタイプの調査において特に大きい。</li> <li>✓複雑な構成や、以前の回答との継続性が求められる調査内容での活用が期待される。</li> </ul>
4)検討の体制	<ul style="list-style-type: none"> <li>－利用可能な技術を都度評価できるように長期的な検討及び試行－</li> <li>✓技術の進歩に伴って、統計調査に活用可能な技術が増える中、今後も新たな技術が生まれる可能性が高い。</li> <li>✓継続的な技術進歩に対応できるよう、技術を都度評価しながらの試行が求められる。</li> </ul>
2. 結果活用ステージにおける示唆・課題	
1)情報通信技術を活用した情報提供内容	<ul style="list-style-type: none"> <li>－利用者の利用目的・レベルに応じた情報提供－</li> <li>✓集計表やマイクロデータなど、統計調査の結果について様々な情報提供が行われる中で、情報通信技術を活用することで、利用者の利用目的やレベル(結果概要の把握や複雑な分析など)に対して、より沿った形での提供が可能となる。</li> </ul>
2)情報通信技術を活用した情報の提供方法	<ul style="list-style-type: none"> <li>－利用者の利用シーンを想定した支援－</li> <li>✓データ提供のみにとどまらず、分析に必要となるソフトウェアやサービスを紹介・提供するなど、付随して必要な機能についても提供することで、利活用をより促進することが求められる。</li> </ul>

次ページ以降で、上記の示唆・課題のそれぞれについて、海外における実態を踏まえた整理を行っている。本調査研究で対象とした海外事例は、アメリカ及びカナダにおいて情報通信技術が活用されている統計調査あるいは統計機関(研究機関を含む)の一部である。

## 第5章 統計の作成、提供等において情報通信技術を活用する際の示唆・課題

### 1. 調査実施ステージにおける示唆・課題

調査実施ステージにおいては、「情報通信技術の位置付け」、「情報通信技術の活用目的」、「統計調査の回答者、回答内容を踏まえた活用」及び「検討の体制」の4点に関する示唆・課題が指摘される。

#### 1) 情報通信技術の位置付け

調査実施ステージにおいて情報通信技術を活用した取組を行っている海外事例を見ると、CATI や CAPI、インターネット回答といった、情報通信技術を活用した回答方法は、複数ある回答方法の1つとして位置付けられている。アメリカ労働統計局の CES のように、数十年にわたって実施されてきた統計調査においても、過去の回答方法を引き継ぎながら、情報通信技術を活用した回答方法が導入されており、回答方法について急激な変更を加えずに追加・修正が行われている。伝統的な紙による調査票を用いた郵送調査や、調査員による聞き取りも併用されている中で、情報通信技術を活用した回答方法も用意することで、後述の「2) 情報通信技術の活用目的」に資するよう配慮がなされている。

また、様々な情報通信技術について、都度評価・検討を加えながら、統計調査への活用が行われている。例えば、下記 CES における取組の歴史を見ると、検討が行われたものの採用されなかった VR(Voice Recognition: 音声認識)では、回答者による電話を用いた自記式という点で、同時期にすでに利用が進んでいた TDE と重複することから採用が見送られている。

－CES における回答方法の変遷－

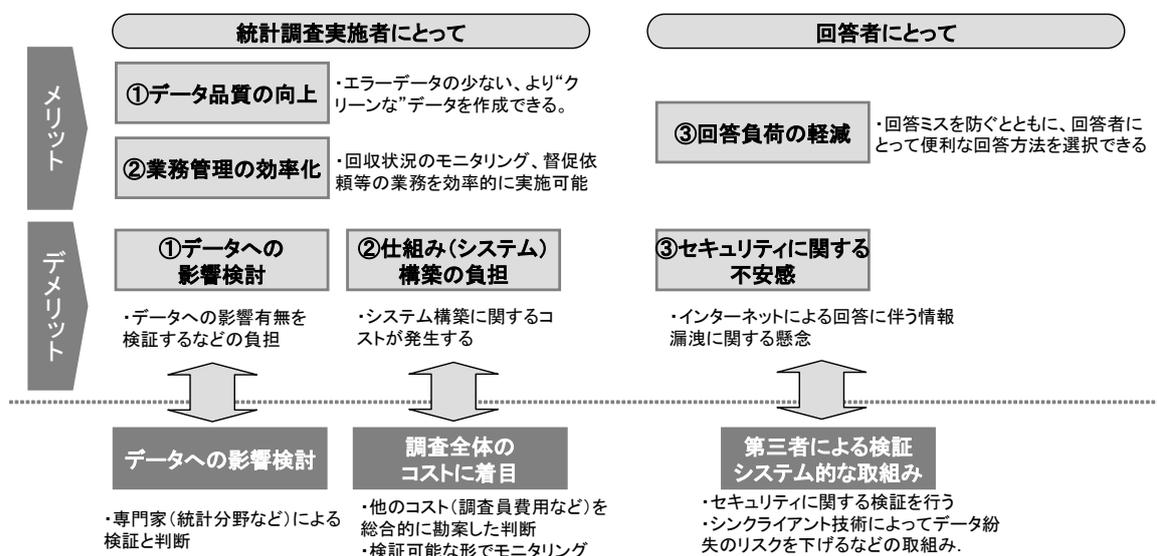
回答方法・導入年	詳細
CATI 1984 年	毎月、調査員が回答者に電話をかけて質問 回答データは同時に入力されるとともに必要に応じて修正される
TDE 1986 年	回答者が通話料無料の番号に電話をかけて、自動化された質問に回答 回答者は電話のキーパッドを使用する
EDI 1994 年	回答者は自企業のシステムからのデータを送付 ファイルは電子的に送信・受信される
FAX 1995 年	毎月、調査票が FAX される 回答者は調査票に記入し FAX で返送する
Web 1996 年	回答者は BLS の Web サイトと接続し回答を入力する 回答終了前にデータは自動的に修正される

このように情報通信技術を活用した回答方法は、複数の異なる回答方法の1つとして位置付けること、また利用可能な技術を都度取り入れながら、より良い活用方法を検討することが求められる。インターネットの普及率や PC リテラシーといった制約条件は存在するものの、海外での事例からは自記入式の回答方法の1つとしてのインターネット回答や、調査員が介在する回答方法の1つとしての CATI といったように、既存の回答方法をより良くするための取組として捉えることが重視されている。日本においても同様の制約条件は存在しており、またユーザーサポートやサーバー負荷への対応といった課題も存在することから、情報通信技術を活用した回答方法のみの実施を推奨するのではなく、これまでの回答方法に追加する形で情報通信技術を活用した回答方法を整えていくことが望ましいと考えられる。

2) 情報通信技術の活用目的

海外事例を見ると、情報通信技術を導入した経緯・目的として、「回答負荷の軽減」、「データ品質の向上」、「コスト削減」の3点が挙げられている。これらのポイントについて、統計実施者及び回答者の視点から、メリット・デメリット及びデメリットへの対応を整理すると以下ようになる。

— 情報通信技術の活用によるメリット・デメリットおよび対応 —



[メリット]

①データ品質の向上

統計調査実施者にとっては、CATI や CAPI、インターネットを活用することで、調査票のロジックに反する回答や未回答を防ぎ、エラーの発生率を減らすことが可能となる。また、毎月調査を行うような内容であれば、前月の回答結果を表示した上で、回答誤りが起こらないように注意喚起を行うなど、様々な工夫を行うことが可能である。それらの工夫によって、回答データの整合性チェックや回答者に対する再確認にかかる時間やコストを節約することが可能となる。特に質問の構造が複雑な調査においては、そのようなデータクリーニング作業を減らすことが、調査全体に係るコストを削減することにもつながる。

②業務管理の効率化

アメリカセンサス局で運用されている Centurion のような共通プラットフォームが整備されると、複数の調査を統合して管理することも可能となる。Centurion では、調査票の作成や回答データの管理を含め、実査時の管理も統合管理が行われており、それによって、回収状況のモニタリングや督促依頼等の業務管理を効率的に行うことが可能となっている。また、調査票(あるいは調査プログラム)の修正変更の履歴管理といった、実査時に限らず、調査設計に関わる業務についても業務の効率化が図られる。

③回答負荷の軽減

回答者にとっては、回答方法が複数提示されることで、個々にとって最も便利な回答方法を選択することができる。同時に、インターネット回答ではエラーメッセージを表示させるなど、複雑な調査内容の調査でも、回答しやすい環境も整えられる。そうした環境が整えられることは、回収率の向上にも寄与し、統計調査の品質の向上にもつながると考えられる。

## 第5章 統計の作成、提供等において情報通信技術を活用する際の示唆・課題

### [デメリット及び対応]

#### ①データへの影響検討

情報通信技術を活用して新たな回答方法を導入するにあたっては、回答方法の変更によるデータへの影響についても検討が必要とされる。海外における事例からは、就労時間や賃金、雇用状態などのファクトデータでは、回答方法の差異による影響は無視できると判断されることが多いが、検討にあたっては以下のような視点から検証が行われていた。

－カナダ統計局 LFS において検証が行われた項目－

EQ 回答者の属性 推計 産業および職業の分布 定性的なフィードバック 回収率 世帯構成員数の変化 補完推計および代入	非回答(スキップ) 就労状態の変化 回答の分散 回答の平均値 失業率の推計
---	---

上記のように、データへの影響検証にあたっては、回答者の属性や推計結果についても検証が行われており、統計学等に精通した専門家との協働も行われている。また、公的統計調査においては時系列での比較可能性が重視される。特に、日本における統計調査の多くが紙による調査を中心に行われている中で、情報通信技術を活用した新たな回答方法を導入する際には、結果の公表にあたって、回答方法別の回答割合を公表するなど、データに影響を及ぼす要因となり得る事項についてもモニタリング・公開することを検討することが望ましいと考えられる。また、同一客体に対して繰り返し調査を行う場合には、最初は調査員が介在した回答方法を通じて調査に習熟した上で、自記入式の回答方法へと移行するなど、可能な限りデータへの影響を小さくするような工夫も求められる。

#### ②仕組み(システム)構築の負担

インターネット回答の仕組み構築や EDI によるデータ交換に係る開発費用など、情報通信技術を活用するにあたっては、仕組みを構築するための負担が発生する。負担が発生する代わりに、情報通信技術の活用によって効率化を図ることとなるが、海外事例では、1サンプルあたりの回収費用の比較を行うなど、その効果をモニタリング可能な形で検証・評価することが意識されており、情報通信技術導入後の改善も意識されている。

また、分散型の統計システムを持つ日本においては、個別の統計調査に特有な事情があるため、様々な統計調査に対応できるよう柔軟性を持った仕組みを構築することが求められる。基幹統計に限っても 50 以上の統計調査がある中で、仕組みを個別に作成するのではなく、共通化したプラットフォームを構築することで負担を分け合うなどの工夫が求められる。現在の e-stat は、日本における統計調査に関する情報について統計利用者にとっての総合的ポータルとしての機能を提供しているが、統計調査の回答者にとってのポータル機能を整備することについても検討を行うことが望ましいと考えられる。

仕組み構築にあたっては、最新の技術動向に精通した民間事業者に委託を行う、あるいは民間事業者と共同での開発を行うなど、外部との連携についても検討を行う必要がある。

### ③セキュリティに関する不安感

回答者側からはインターネット回答に伴う情報漏洩などに対する懸念が生じる場合がある。海外における取組事例では、暗号化された通信環境(SSL 通信)によるデータ保護に加えて、第三者機関等によるチェック体制を整えるなど、回答者からの不安感を減らすための取組みが行われていた。また、CATI や CAPI で使用される端末についても、シンクライアント技術を活用することで、回答データや対象者の情報を端末に保存しないようにするなどの取組みも行われている。

セキュリティを過度に重視すると回収率が低下するなどの問題が生じることも指摘されていた。セキュリティ確保においては、前述のような情報通信技術(SSL 通信など)を活用することで、回答者側の負荷を低く抑えながらセキュリティを確保することが望ましい。日本においては、プライバシーに対する意識が高まる中、より繊細な内容についても回答協力を得られるように、調査員との対面ではない自記式の回答方法を活用することを検討する必要がある。

### ④調査員の負担

オンライン回答方式では、世帯の特定のため回答者が ID 等を入力する必要があるが、回答者が ID 等を入力しなかった場合、世帯の特定にかかる調査員の負担が増える懸念がある。

実際に、東京都でのみオンライン回答方式が導入された平成 22 年国勢調査では、世帯の特定のために回答者が調査区番号に加えて世帯番号を入力する必要があったが、世帯番号の入力漏れが多数発生した。これらの世帯番号が入力されていない調査票は「迷子票」と呼ばれ、調査区と氏名のみから世帯を特定することが必要となり、調査員の負担が増加する結果となった。平成 27 年国勢調査では、ID と世帯番号をあらかじめ紐付け ID のみの入力で見落としを可能にする等の対策が検討されている。

## 3) 統計調査の回答者、回答内容を踏まえた活用

海外における取組事例からは、調査対象者が世帯・個人、企業・事業所のいずれにおいても、情報通信技術の活用が行われている。一方で、調査内容については、回答負荷軽減、データ品質の向上といったメリットは、同一客体に対して定型の内容を繰り返し質問するタイプの調査において特に大きいと考えられる。

客体については、回答者が世帯・個人であれば CATI やインターネット回答が、企業・事業所であれば EDI やインターネット回答が有力な候補となる。一方で、従来の紙による回答に慣れている回答者への配慮といった工夫も行うことが必要とされる。

回答内容については、主として回答者の意識を把握するための統計調査よりも、社会経済の実態を把握するための統計調査全般に対する活用がよりふさわしいと考えられる。今回調査対象となった海外事例では、賃金や労働時間、就業状態といったファクトデータを中心とする統計調査を取上げたが、社会調査の分野において意識データを聴取する調査においても、CATI や CAPI、インターネット回答など複数の回答方法を用意することが一般的であるとの意見も聞かれた。日本における統計調査の中でも、上記のようなファクトデータを継続的に収集するような調査に対して、情報通信技術を活用することが望ましいと考えられる。

## 4) 検討の体制

今回の調査対象となった海外事例では、インハウスで検討されることが多いが、技術の進歩に応じて、その時々利用可能な情報通信技術の評価しながら検討が行われることが指摘されている。海外事例では、回答方法の変更に伴って、調査画面のテストを通じたユーザビリティの確認を行うなど、実導入を行う前にパイロットテストを行うことの重要性が聞かれた。技術面に精通した専門家等の支援も行われており、技術の進歩に伴って統計調査に活用可能な技術が増える中、今後も新たな技術が生まれる可能性が高い。国勢調査において有識者会議を設置し外部専門家の意見を積極的に取り入れる体制を取っているように、継続的な技術進歩に対応できるよう、技術を都度評価しながらの試行が可能となるような体制が必要と考えられる。

## 第5章 統計の作成、提供等において情報通信技術を活用する際の示唆・課題

### 2. 結果活用ステージにおける示唆・課題

結果活用ステージにおいては、「情報通信技術を活用した情報提供内容」及び「情報通信技術を活用した情報の提供方法」に関する示唆・課題が指摘される。

#### 1) 情報通信技術を活用した情報提供内容

集計表やマイクロデータなど、統計調査の結果について様々な情報提供が行われる中で、情報通信技術を活用することで、利用者の利用目的やレベル(結果概要の把握や複雑な分析など)に対して、より沿った形での提供が可能となる。

例えば、人の目による最終的なチェックは必要ではあるものの、システムを活用して(必要最低限の)セキュリティ確保を行うことで、オンライン上でマイクロデータにアクセスして分析することが可能となっている。また、GIS情報と組み合わせた形での情報提供も可能となるなど、より多様な情報提供が可能となっている。海外事例を見ると、セキュリティを確保しつつも、より多様な情報提供を可能とすべく、商用パッケージソフトウェアなど既存の利用可能な技術が活用されている。

#### 2) 情報通信技術を活用した情報の提供方法

データ提供のみにとどまらず、分析に必要となるソフトウェアやサービスを紹介・提供するなど、付随して必要な機能についても提供することで、データの利活用をより促進することが可能となる。海外事例では、利用者側のニーズ(高度な分析～簡易な集計表、など)に応じた情報を提供するため、ツール(ソフトウェア)の整備を含めた利用しやすい環境整備が進められている。また、利用者がいつでも・どこからでもアクセスできる環境に対して必要性が高まる中で、リモートアクセス技術(サーバー仮想化)の活用も進められている。特に、リモートアクセス技術を利用することで、オンライン上で複雑な分析が可能となるなど、より高度な利用を求める利用者にとって、より利便性が高まることにつながる。

上記2つに共通する課題として、システムリソースに関する課題が挙げられる。データ利用の活性化は同時にシステム負荷増加につながるため、必要となるリソースを慎重に検討する必要がある。

ーリモートアクセス技術の例ー

VMWare	Citrix
-ICPSRにて利用されている -個別ユーザーに対するリソースの割当てを行う -ユーザーが多数ログインしていてもパフォーマンスに影響しないものの、割当て以上のパフォーマンスを出すことはできない	-NORCの dataenclaveにて利用されている -システム全体のリソースを個別ユーザーが利用できる -ユーザーが多数ログインするとパフォーマンスが低下するものの、ログイン数が少なければパフォーマンスが向上する

上記のようなリモートアクセス技術の例に記載されるように、結果活用の中でもどのような機能を提供するかの優先度付け・見極めが重要となる。特に、定期的に実施される統計調査の場合、追加されるデータ量に対応する必要があるため、どのような技術を活用するかについても慎重な検討が必要とされる。

## 附属資料 目次

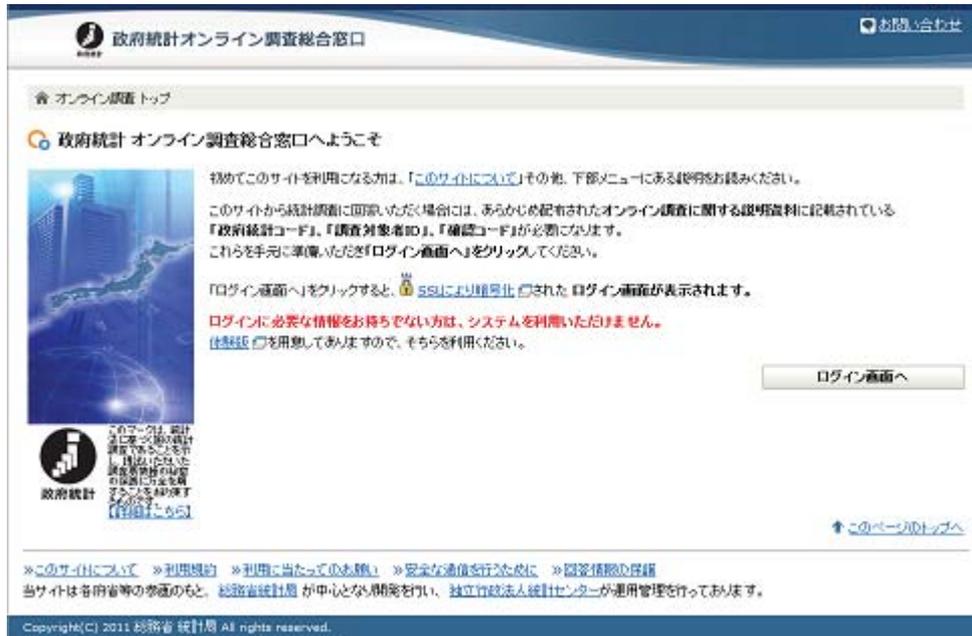
<b>I.国内における情報通信技術の活用に関する取組</b>	<b>附属 1</b>
1 統計局	附属 1
1) 調査票例	
2) オンライン調査システムについて	
2 統計センター	附属 12
1) 統計データの二次的利用に関する統計センターの取組状況	
2) 統計におけるオープンデータの高度化	
<b>II.海外における情報通信技術の活用に関する取組</b>	<b>附属 35</b>
1 アメリカセンサス局(United States Census Bureau)	附属35
・ Centurion 概要	
2 アメリカ労働統計局(United States Bureau of Labor Statistics)	附属49
1) CES 調査票例	
2)IDCF(Internet Data Collection Facility)概要	
3 カナダ統計局(Statistics Canada)	附属63
1) LFS におけるインターネット回答の検証について	
2) マイクロデータへのアクセスについて	
<b>III.参考ウェブサイト・文献一覧</b>	<b>附属110</b>





2) オンライン調査システムについて

初期アクセス画面



ログイン画面



出所) 統計局 政府統計オンライン調査総合窓口ウェブサイト

確認コードの変更画面(最初のアクセス時のみ)

政府統計オンライン調査総合窓口

お問い合わせ ヘルプ ログアウト

ホーム > オンライン調査トップ > ログイン > 確認コードの変更

### 確認コードの変更

確認コードについて : 確認コードは、調査対象者IDが統計調査の対象の方に利用されていることを確認するためのもので、暗証番号やパスワードと同様の役割があります。

初めてのログイン : あらかじめ配布された確認コードは調査に関係しない第三者の目に触れる可能性があるため、初回ログイン時は確認コードを変更していただく必要があります。

2度目以降のログイン : 確認コードは安全のため定期的な更新をお勧めします。  
このシステムでは統計調査ごとに確認コードの有効期間を定めており、一定期間を経過すると確認コードの変更を求められます。  
不明な点についてはお上の「お問い合わせ」から表示される問い合わせ先にご連絡ください。

#### 確認コード情報

新しい確認コード(10文字以上)に変更してください。  
確認コードは、大文字、小文字、数字をそれぞれ1文字以上含む文字列としてください。  
セキュリティを高めるため辞書攻撃を防止する対策を行っておりますので、推測されやすい文字列は利用しないでください。  
推測されやすい文字列例⇒ Password1、Japan123、Windows1等  
入力禁止のため同じ確認コードを2回連続して入力できません。  
変更した確認コードを忘れてしまった場合には、お上の「お問い合わせ」から表示される問い合わせ先にご連絡ください。

確認コード(必須)	<input type="text"/>	(半角英数8文字以上32文字以内)
確認コード(確認用)(必須)	<input type="text"/>	(半角英数8文字以上32文字以内)

変更

[このページのトップへ](#)

※このサイトについて ※利用規約 ※利用にあたってのお願い ※安全な通信を行うために ※調査情報の保護  
当サイトは各府省等の参画のもと、総務省統計局が中心となり開発を行い、独立行政法人統計センターが運用管理を行っております。

Copyright(C) 2011 総務省 統計局 All rights reserved.

連絡先情報の登録画面(最初のアクセス時のみ)

政府統計オンライン調査総合窓口

お問い合わせ ヘルプ ログアウト

ホーム > オンライン調査トップ > ログイン > 連絡先情報の登録

### 連絡先情報の登録

#### 連絡先情報

連絡先情報を入力後、「登録」ボタンをクリックしてください。  
ここで登録されたメールアドレスは、調査票の受付状況メールの送信など各種への連絡に使用します。

メールアドレス	<input type="text"/>	(半角英数60文字以内)
---------	----------------------	--------------

登録

[このページのトップへ](#)

※このサイトについて ※利用規約 ※利用にあたってのお願い ※安全な通信を行うために ※調査情報の保護  
当サイトは各府省等の参画のもと、総務省統計局が中心となり開発を行い、独立行政法人統計センターが運用管理を行っております。

Copyright(C) 2011 総務省 統計局 All rights reserved.

## 連絡先情報の確認画面

政府統計オンライン調査総合窓口
お問い合わせ | ヘルプ | ログアウト

[ホーム](#) > [オンライン調査トップ](#) > [ログイン](#) > [連絡先情報の確認](#)

### 連絡先情報の確認

登録いただいた連絡先情報は以下のとおりです。  
 確認いただき、よろしければ、「調査票の一覧へ」ボタンをクリックしてください。  
 表示内容に変更がある場合は、「連絡先変更へ」ボタンをクリックしてください。

メールアドレス	abc@xxx.ne.jp
---------	---------------

調査票の一覧へ
連絡先変更へ

[↑このページのトップへ](#)

※このサイトについて   ※利用規約   ※利用にあたってのお願い   ※安全な通信を行うために   ※回答情報の保護  
 当サイトは各府省等の参画のもと、総務省統計局が中心となり開発を行い、独立行政法人統計センターが運用管理を行っております。

Copyright(C) 2011 総務省 統計局 All rights reserved.

## 調査票の一覧画面

政府統計オンライン調査総合窓口
お問い合わせ | ヘルプ | ログアウト

[ホーム](#) > [オンライン調査トップ](#) > [ログイン](#) > [調査票の一覧](#)

### 調査票の一覧

- 回答する調査票はあらかじめ配布された説明資料で確認し、調査票一覧から選択し選択欄にチェックマークをつける。調査票が複数ある場合は、必要なものを選択する。 「調査票の取得」ボタンをクリックしてください。調査票のダウンロードが開始されます。ただし、HTML形式の調査票は他の調査票と同時にダウンロードすることができません。
- ダウンロードした調査票を保存する場合は、必ずファイル名を変更して保存してください。
- 「記入例」欄に「表示」がある調査票には、記入例が掲載されています。記入例を見る場合は該当する行の「表示」をクリックし、記入例をダウンロードしてください。
- 調査票一覧の「受付番号」、「回答日」に値が表示されている調査票は、既に回答いただいた調査票です。
- 「回答確認」欄に「確認」がある調査票は、回答いただいた内容を確認することができます。回答を確認する場合は該当する行の「確認」をクリックし、回答が入力された電子調査票をダウンロードしてください。
- その他、「確認コード・連絡先情報を変更」、「複数調査票を一括して回答する」、「複数統計調査の調査対象者IDをまとめて管理する」等が必要であれば下部のリンクから行ってください。
- なお、回答をお断りしている調査票とは別に、アンケートを任意でお断りしている場合があります。「アンケートに回答」ボタンが表示されている場合には、アンケートにも協力をお断りいただけます。

確認コード・連絡先情報を変更したい場合   : [確認コード・連絡先情報の変更](#)  
 複数調査票の回答データをまとめて回答したい場合   : [回答の一括取得](#)  
 複数統計調査の調査対象者IDを統合したい場合   : [統合処理の追加](#)

調査票の取得

#### オンライン調査テスト3 対象者特定済

回答する調査票を選択し、「調査票の取得」ボタンをクリックしてください。

選択	実施時期	電子調査票	ファイル形式	提出期限	受付番号	回答日	記入例	回答確認
<input type="checkbox"/>	平成24年	●●調査票		2012/10/15			<a href="#">表示</a>	<input type="checkbox"/>

アンケートに回答

[↑このページのトップへ](#)

※このサイトについて   ※利用規約   ※利用にあたってのお願い   ※安全な通信を行うために   ※回答情報の保護  
 当サイトは各府省等の参画のもと、総務省統計局が中心となり開発を行い、独立行政法人統計センターが運用管理を行っております。

Copyright(C) 2011 総務省 統計局 All rights reserved.

## 調査票のパスワード設定画面(統計調査によって異なる)

政府統計オンライン調査総合窓口

オンライン調査 トップ > ログイン > 調査票の一覧 > 調査票のパスワード設定

### 調査票のパスワード設定

PDF形式のファイルで作成された調査票の特徴は以下のとおりです。

- 目次前、目次途中の調査票をPCに保存することが可能です。(保存しなければ入力した内容は破棄されます。)
- 保存した調査票はAdobe Readerで開くことにより回答を再開することが可能です。
- 調査票にある「送信」ボタンをクリックすることにより、システムに回答データを送信することができます。
- 送信の際、現在設定されている「確認コード」の入力が必要となります。また、調査票にパスワードを設定した場合には、Adobe Readerで開く際このパスワードの入力が必要となります。

#### 調査票

この調査票は、PDF形式のファイルで作成されており、入力された内容等を他人に見られないようにパスワードを設定することができます。ここで設定するパスワードは、ログイン時に入力していたID及び「確認コード」とは一切関係なく設定することが可能です。パスワードを設定すると、ダウンロードした調査票を印刷の際は、設定したパスワードが必要となりますので、設定したパスワードはメモするなど大切に保管してください。他の調査票も同じパスワードを設定する場合は、「全て同じパスワードを設定」にチェックを入れることでパスワードを一括で設定できます。パスワードの設定も必要としない場合は、空欄のまま「ダウンロード」ボタンをクリックしてください。「ダウンロード」ボタンをクリックすると、調査票のダウンロードが始まります。

パスワード	(半角英数字32文字以内)
パスワード(確認用)	(半角英数字32文字以内)

ダウンロード

このページのトップへ

このサイトについて ※利用規約 ※利用に当たってのお願い ※安全な通信を行うために ※調査情報保護

当サイトは各府省等の奉還のもと、総務省統計局が中心となり開発を行い、独立行政法人統計センターが運用管理を行っております。

Copyright(C) 2012 総務省 統計局 All rights reserved.

## 電子調査票画面(PDF形式)

https://www.e-survey.go.jp/G004020201.a.do - Windows Internet Explorer

https://www.e-survey.go.jp/G004020201.a.do

https://www.e-survey.go.jp/G004030301.a.do

このコンテンツはフォームモードで表示されています。このフォームに入力したデータは保存できません。記録を残したい場合は、入力済みのフォームを印刷してください。

指定統計第7号 平成 19 年 10 月 1 日

就業構造基本調査調査票

15歳以上の個人について記入してください

総務省統計局

クリア

1 氏名・男女の別及び配偶者(妻又は夫)の有無

2 婚姻上の続柄

3 出生の年月

4 教育

5 1年動向

6 あなたはるだん何か収入になる仕事をしていますか

完了

印刷ボタン

電子調査票画面(HTML形式)

指定統計第87号 平成19年10月1日  
**就業構造基本調査調査票**  
 ◎15歳以上の各人について記入してください

1 氏名・男女の別及び配偶者（妻又は夫）の有無  
 ● 配偶者の有無は届出の有無に準拠なく記入してください

2 世帯主との続き柄  
 ● 世帯主の配偶者の祖父母・兄弟姉妹はそれぞれ「祖父・祖母」・「兄弟姉妹」に含めます

3 出生の年月  
 ● 該当する番号又は西暦について記入した上で年及び月を記入してください

4 教育  
 ● 現在 学校に在学しているかどうかについて記入した上で、空白に就いて記入してください  
 ● 「卒業」の人は最終卒業学校（中・高等学校またはその相当の学校）について記入してください  
 ● 予備校などはここでいう学校には含めません

5 1年前にどこに住んでいましたか  
 ● 東京都府県及び特別区等市町村の区分を記入してください

調査票回答の受付状況画面

政府統計オンライン調査総合窓口

ホーム > オンライン調査 トップ > ログイン > 調査票の一覧 > 調査票回答の受付確認

**調査票回答の受付状況**

調査票の受付状況

統計調査名	●就業
実施時期	平成24年度
調査票名	●就業票
調査対象者ID	HDZED11
キー項目	48123400000011
受付番号	078KH4153001
受付結果	調査票回答を登録しました。ご回答ありがとうございました。

調査票の一覧へ アンケートに回答

●「政府統計オンライン調査総合窓口」での回答状況は、上記「受付結果」のとおりとなっております。メールアドレスを登録していた方には、受付状況をお知らせするメールも送信しています。回答状況を確認する必要が生じた場合には、再度ログインし「調査票の一覧」画面で確認することができます。通信上のトラブル等で受付が正常に行われなかった場合には、随時判別のため右上の「ヘルプ」をご利用ください。また、操作が繰り返しても受付が正常に行われなかった場合には、右上の「お問い合わせ」で表示されるお問い合わせ先にご相談ください。

● 受付番号は調査に回答していただいたこととなります。（受付状況をお知らせするメールにも記載されています。）

● 「アンケートに回答」ボタンが表示されている場合は、アンケートにも協力をお願いします。

● 引き続き同じ統計調査の別の調査票に回答いただく場合には、「調査票の一覧へ」ボタンをクリックしてください。それ以外の場合は、右上の「ログアウト」をクリックしてください。

このページのトップへ

このサイトについて >> 利用規約 >> 利用にあたってのお願い >> 安全な通信を行うために >> 個人情報保護  
 当サイトは各府省等の委託のもと、総務省統計局が中心となり開発を行い、独立行政法人統計センターが運用管理を行っております。

Copyright(C) 2012 総務省 統計局 All rights reserved.

調査票の一覧画面(受付後)

政府統計オンライン調査総合窓口

ホーム > オンライン調査トップ > ログイン > 調査票の一覧

### 調査票の一覧

- 回答する調査票はあらかじめ配布された説明資料で確認し、調査票一覧から選択(選択欄にチェックマークをつける。調査票が複数ある場合は、必要なものを選択)して、「調査票の取得」ボタンをクリックしてください。調査票のダウンロードが滞ります。ただし、HTML形式の調査票は他の調査票と同時にダウンロードすることができません。
- ダウンロードした調査票を保存する場合は、必ずファイル名を変更して保存してください。
- 「記入例」欄に「表示」がある調査票には、記入例が表示されています。記入例を見る場合は該当する行の「表示」をクリック、記入例をダウンロードしてください。
- 調査票一覧の「受付番号」、「回答日」に値が表示されている調査票は、既に回答済みの調査票です。
- 「回答確認」欄に「確認」がある調査票は、回答したい内容を確認することができます。回答を確認する場合は該当する行の「確認」をクリック、回答が入力された電子調査票をダウンロードしてください。
- その他、「確認コード・連絡先情報を変更」、「複数調査票を一括して回答する」、「複数統計調査の調査対象者IDをまとめて管理する」等が必要であれば下部のリンクから行ってください。
- なお、回答をお断りしている調査票とは別に、アンケートを任意でお断りしている場合があります。「アンケートに回答」ボタンが表示されている場合には、アンケートにも協力をお断りいただけます。

確認コード・連絡先情報を変更したい場合 : [確認コード・連絡先情報の変更](#)  
 複数調査票の回答データをまとめて回答したい場合 : [回答の一括登録](#)  
 複数統計調査の調査対象者IDを統合したい場合 : [統合IDへ追加](#)

**調査票の取得**

調査

回答する調査票を選択し、「調査票の取得」ボタンをクリックしてください。

選択	実施時期	電子調査票	ファイル形式	提出期限	受付番号	回答日	記入例	回答確認
<input type="checkbox"/>	平成24年度	● 調査票		2012/08/08	078KH4153001	2007/08/20	表示	<input type="checkbox"/>

[アンケートに回答](#)

[このページトップへ](#)

※このサイトについて ※利用規約 ※利用に当たってのお願い ※安全な通信を行うために ※調査情報の保護  
 当サイトは各府省等の委託のもと、総務省統計局が中心となり開発を行い、独立行政法人統計センターが運用管理を行っております。

Copyright(C) 2011 総務省 統計局 All rights reserved.

ブラウザ画面で開いた電子調査票(PDF形式)

https://www.e-survey.go.jp/G004020201.a.do - Windows Internet Explorer

https://www.e-survey.go.jp/G004020201.a.do

https://www.e-survey.go.jp/G004030301.a.do

このドキュメントはフォームフィールドが埋まっています。このフォームに入力したデータは保存できません。応答を修正したい場合は、入力済みのフォームを印刷してください。

クリア

指定統計番号 平成 23 年 10 月 1 日  
**就業構造基本調査調査票**  
 ◎15歳以上の本人について記入してください

1 氏名・男女の別及び配偶者(妻又は夫)の有無

2 世帯主との続柄

3 出生の年月

4 教育

5 1年前にはどこに住んでいましたか

6 あなたはだん何か収入になる仕事をしていますか

不規則ページ

ブラウザ画面で開いた電子調査票 (HTML 形式)

指定統計第67号 平成 19 年 10 月 1 日  
就業構造基本調査調査票  
※15歳以上の各人について記入してください

1 氏名・男女の別及び配偶者（妻又は夫）の有無  
・配偶者の有無は届出の有無に限らず記入してください

2 世帯主との続き柄  
・世帯主の配偶者の祖父・兄弟姉妹はそれぞれ「祖父」・「兄弟姉妹」に当てます

3 出生の年月  
・該当する番号又は西暦について記入した上で年及び月を記入してください

4 教育  
・現在 学校に在学しているかどうかについて記入した上で 空白に就いて記入してください  
・「卒業」の人は最終卒業学校（中途退学した人はその時の学校）について記入してください  
・予備校などはここでいう学校には含めません

5 1年前にどこに住んでいましたか  
・東京都内及び指定都市以外の場合は「東京都区部及び指定都市の別の市区町村」を選択してください

Acrobat Reader 単独で開いた電子調査票 (PDF 形式)

指定統計第67号 平成 19 年 10 月 1 日  
就業構造基本調査調査票  
※15歳以上の各人について記入してください

1 氏名・男女の別及び配偶者（妻又は夫）の有無  
・配偶者の有無は届出の有無に限らず記入してください

2 世帯主との続き柄  
・世帯主の配偶者の祖父・兄弟姉妹はそれぞれ「祖父」・「兄弟姉妹」に当てます

3 出生の年月  
・該当する番号又は西暦について記入した上で年及び月を記入してください

4 教育  
・現在 学校に在学しているかどうかについて記入した上で 空白に就いて記入してください  
・「卒業」の人は最終卒業学校（中途退学した人はその時の学校）について記入してください  
・予備校などはここでいう学校には含めません

5 1年前にどこに住んでいましたか  
・東京都内及び指定都市以外の場合は「東京都区部及び指定都市の別の市区町村」を選択してください

6 あなたはどんな収入になる仕事をしていますか  
・ここで「仕事をしている」とは部分時間労働者も含みます  
・収入になる仕事には 家族（養育を含む）の労働者も含まれます  
・家事には 育児・介護・看護などを含みます  
・遠方では「1」を、職業上の収入にのみならず「専業主婦・専業主夫」の収入も含みます

### ログイン情報の確認画面

JavaScript ウィンドウ

政府統計コード: 4ABK

調査対象者ID: abc0001

確認コード:

警告: JavaScript ウィンドウ

### セキュリティ警告画面

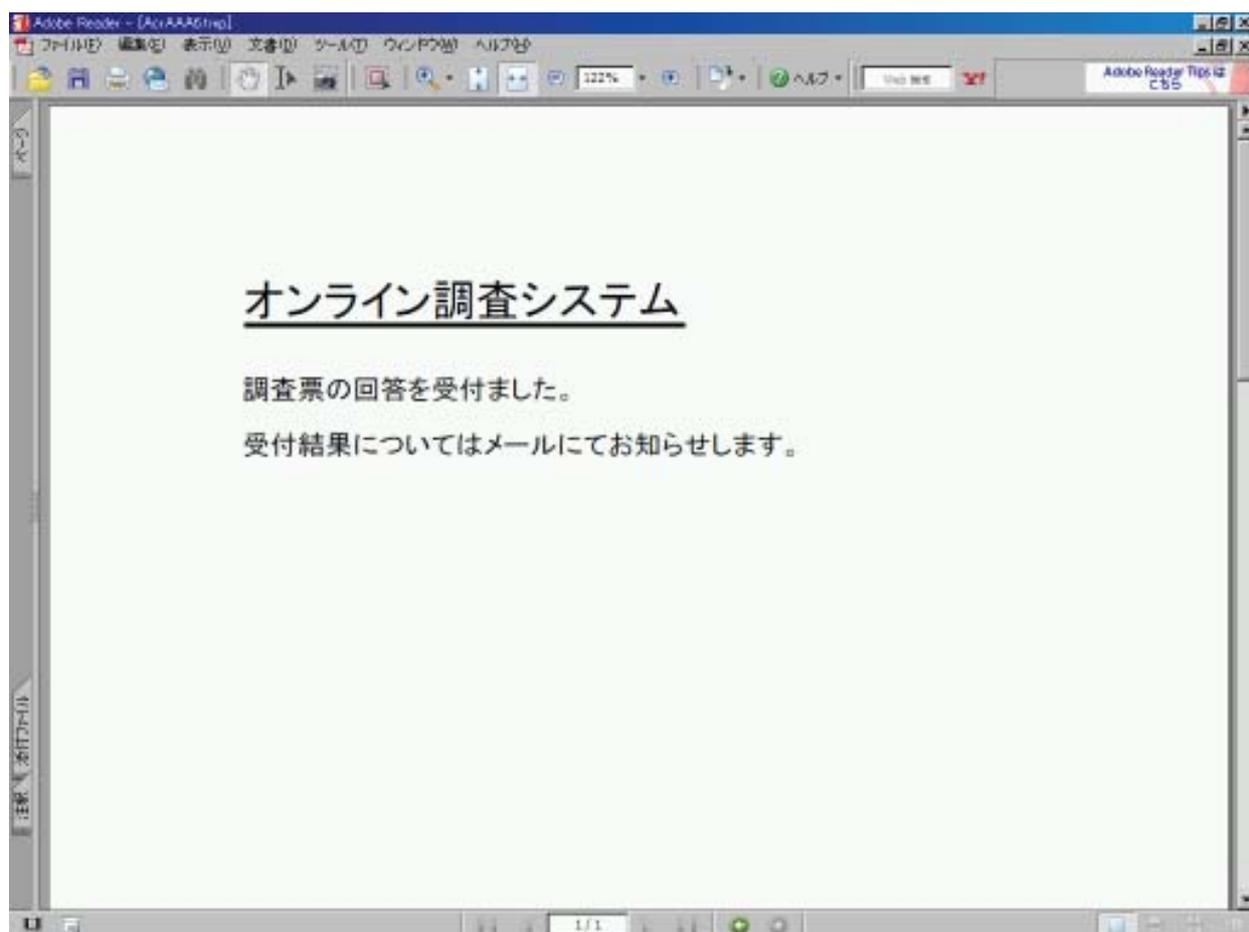
セキュリティ警告

 文書は次のサイトに接続しようとしています:  
[https://www.e-survey.go.jp/top2/GD08010101a.do?first\\_request=on](https://www.e-survey.go.jp/top2/GD08010101a.do?first_request=on)

このサイトを信頼する場合は「許可」をクリックします。そうでない場合は「ブロック」をクリックします。

このサイトに対するアクションを保存する: [www.e-survey.go.jp](http://www.e-survey.go.jp)

調査票回答の受付状況画面(Adobe Reader 単独で開いた電子調査票の送信)



統計センター

1) 統計データの二次的利用に関する統計センターの取組状況

# 統計データの二次的利用に 関する 統計センターの取組状況

2013年度 統計関連学会連合大会

平成25年9月9日

独立行政法人統計センター 亀本 信康

齋藤 敦



出所) 統計センター ウェブサイト

## <報告内容>

---

- 1 公的統計の二次的利用のための枠組について
- 2 統計センターにおける  
匿名データ及びオーダーメイド集計の提供について
- 3 擬似マイクロデータの試行的提供について
- 4 統計センターと学術研究機関等との連携について

## 1 公的統計の二次的利用のための枠組について

### 公的統計とは？

国などの行政機関等が実施している統計調査  
例: 国勢調査(総務省)、労働力調査(総務省)など

### <統計法の改正>

#### 統計法(平成19年法律第53号)

- 平成19年5月 公布
- 平成21年4月 全面施行

#### 【基本理念】

「行政のための統計」から「**社会の情報基盤としての統計**」へ

- ① 公的統計の体系的・計画的整備の推進
- ② **統計データの有効利用の促進**
- ③ 統計調査の対象者の秘密保護の強化
- ④ 統計委員会の設置

マイクロデータの二次的利用  
・匿名データの提供  
・委託による統計の作成等  
(オーダーメイド集計)

## 2 統計センターにおける匿名データ及びオーダーメイド集計の提供

### 匿名データの提供

一般からの申出を受け、利用要件を満たした申出者に対し、特定の個人又は団体等が識別できないように加工して作成した調査票情報の利用を一定期間認める制度。匿名データを利用することで、行政機関が作成していない統計表の作成のみならず、多変量解析など**マイクロデータ**に基づく実証分析を行うことが可能。

#### 【利用要件】

- ・統計の作成または統計的研究にのみ利用されること
- ・学術研究目的、高等教育目的等の用に供することを直接の目的とすること
- ・学術研究の成果、高等教育等の内容等が公表され、社会に還元されること
- ・匿名データが適切に管理されること

#### 統計センターが各府省からの委託を受けて提供している統計調査(平成25年8月末現在)

- |                                 |                               |  |
|---------------------------------|-------------------------------|--|
| ■ 全国消費実態調査<br>(平成元年、6年、11年、16年) | ■ 住宅・土地統計調査<br>(平成5年、10年、15年) | ■ 社会生活基本調査<br>【生活時間編・生活行動編】<br>(平成3年、8年、13年、18年) |
| ■ 就業構造基本調査<br>(平成4年、9年、14年)     | ■ 労働力調査<br>(平成元年1月～平成20年12月)  |  |

25年度 提供予定

- |                       |                          |
|-----------------------|--------------------------|
| ■ 国勢調査<br>(平成12年、17年) | ■ 労働力調査<br>(平成21年1月～12月) |
|-----------------------|--------------------------|

～ 利用可能な統計調査は、今後、さらに拡大予定 ～

(注)括弧内の年次は、提供している調査年次を示す。

## (参考) 匿名データの提供の流れ



匿名データの利用に関するFAQ: <http://www.nstac.go.jp/services/faq-anonymity.html>

## (参考) 匿名データの利用実績例

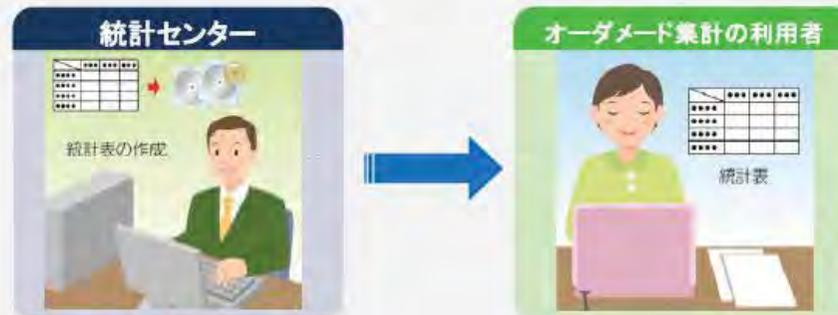
利用目的	調査名	研究の名称	
学 術 研 究 目 的	社会生活基本調査	正規雇用者における平日の労働時間と休息時間 —「社会生活基本調査」マイクロデータによる分析— 生活行動からみる高齢者の行動特性について —社会生活基本調査の匿名データを用いて—	
		子供のいる世帯における夫と妻の2次活動時間の差異について —社会生活基本調査の匿名データを用いて—	
		趣味・娯楽活動の時間について 個人・世帯属性と行楽・観光旅行行動の関係	
		全国消費実態調査	等価尺度の推計と比較—消費上の尺度・制度的尺度・OECD尺度— 『季刊社会保障研究』Vol.48 Spring 2013 No.4 所得格差変動の年齢階級別要因分解 :全国消費実態調査マイクロデータを用いて
		全国消費実態調査 住宅・土地統計調査	持家取得における既婚女性の就業の役割
	就業構造基本調査	転職経験および転職理由と転職希望意識との関連について —就業構造基本調査匿名データによる統計分析— 若年者就業率における賃金弾力性の推定 女性事務職の賃金と就業行動 —男女雇用機会均等法施行後の三時点比較— 税負担と労働供給 —「日本労働研究雑誌」No.605 2010.12— 若者の有業・無業状態における属性の考察 女性の働き方と少子化に関する考察	
	高 等 教 育 目 的	就業構造基本調査	一橋大学大学院経済学研究科「演習」(労働経済学Ⅰ)
		全国消費実態調査	
		社会生活基本調査	

論文等: <http://www.nstac.go.jp/services/jisseki.html>

## 2 統計センターにおける匿名データ及びオーダーメイド集計の提供について

### オーダーメイド集計

一般からの委託を受けて、利用要件を満たした申出者に対し、調査票情報を用いて集計を行い、その結果の提供を行う制度。オーダーメイド集計を利用することで、行政機関等が作成していない統計表に基づいた分析が可能。



#### 【利用要件】

- ・統計の作成または統計的研究にのみ利用されること
- ・学術研究目的または高等教育目的の用に供することを直接の目的とすること
- ・学術研究の成果または高等教育の内容が公表され、社会に還元されること

# オーダーメイド集計

統計センターが各府省からの委託を受けて提供している統計調査(平成25年9月2日現在)

総務省

- 国勢調査 (昭和55、60年、平成2年、7年、12年、17年)
- 住宅・土地統計調査 (昭和63年、平成5年、10年、15年、20年)
- 就業構造基本調査 (昭和62年、平成4年、9年、14年、19年)
- 社会生活基本調査
  - 【生活行動編】(昭和61年、平成3年、8年、13年、18年、23年)
  - 【生活時間編】(平成3年、8年、13年、18年、23年)
- 労働力調査
  - 【基礎調査票】(昭和55年1月～平成23年12月)
  - 【特定調査票】(平成14年1月～平成23年12月)
- 家計調査 (平成元年1月～平成24年12月)
- 全国消費実態調査
  - 【家計収支編、品目編】(平成18年、21年)
  - 【家計資産編(純資産・総資産)、家計収支編曜日別、品目編購入先別】(平成16年)
- 家計消費状況調査 (平成14年1月～平成24年12月)
- 家計消費実態調査
  - 【家計資産編(純資産・総資産)、家計収支編曜日別、品目編購入先・購入地域別】(平成21年)

内閣府

文部科学省

厚生労働省

国土交通省

- 消費動向調査 (平成19年4月～25年3月)
- 学校基本調査
  - 【高等教育機関編】(平成20～23年度)
  - 【初等教育機関編】(平成20～22年度)
- 賃金構造基本統計調査 (平成18年～24年)
- 建築着工統計調査 (平成21年4月～25年3月)
- 企業行動に関するアンケート調査 (平成18年度～24年度)

25年度 提供予定

- 労働力調査 (平成24年1月～12月)
  - 【基礎調査票】、【特定調査票】
- 学校基本調査
  - 【高等教育機関編】(平成24年度)

(注)括弧内の年次は、提供している調査年次を示す。

～ 利用可能な統計調査は、今後、さらに拡大予定 ～

## (参考) オーダーメイド集計の流れ



## (参考) オーダーメイド集計の利用実績例

調査名	研究の名称
国勢調査	岡山大学大学院社会文化科学研究科紀要第35号(2013.3)
	・2005年国勢調査にみる在日外国人女性の結婚と仕事・住居
	・2005年国勢調査にみる外国人の教育
	・2005年国勢調査にみる在日外国人の仕事
	Access to childcare and the employment of women with preschool-aged children in Tokyo
	看護人材の就業率の推移
	―再検討した潜在者数推計方法による結果から―
	地域別経済指標に基づくSDモデルの開発
	在日外国人の仕事
	―2000年国勢調査データの分析から―
全国消費実態調査	家族・ジェンダーからみる在日外国人
	―2000年国勢調査データの分析から―
就業構造基本調査	在学率と通学率から見る在日外国人青少年の教育
	―2000年国勢調査データの分析から―
	エリア・サンプリングにおける問題点
新・家計消費論	―高齢層が支える都市部消費―
近年における都道府県別貧困率の推移について	―ワーキングプアを中心に

論文等：<http://www.nstac.go.jp/services/jisseki.html>

## 2 統計センターにおける匿名データ及びオーダーメイド集計の提供について

### 匿名データ及びオーダーメイド集計の手数料

#### ○匿名データの提供

- ① 依頼1件につき**1,850円**
- ② 匿名データファイル1ファイルにつき**8,500円**
- ③ 次に定める提供用媒体の額
  - CD-R 1枚につき**100円**
  - DVD-R 1枚につき**120円**
- ④ **送付に要する費用**

#### ○委託による統計の作成等（オーダーメイド集計）

- ① 統計の作成等に要する時間1時間までごとに**5,900円**
- ② **統計成果物を郵送で受け取る場合**
  - ア) 次に定める提供用媒体の額
    - CD-R 1枚につき**100円**
    - DVD-R 1枚につき**120円**
  - イ) 送付に要する費用

※ 統計成果物を統計センターからメールで受け取る場合は、媒体の費用や送料は掛かりません。

#### 受益者負担

利用者は手数料を支払う

### 3 擬似マイクロデータの試行的提供

## 擬似マイクロデータの試行的提供

マイクロデータ利用者の裾野を広げ、公的統計の二次的利用の拡大を図るため、試行版の擬似マイクロデータを作成し、試行的提供を行っている。(無償)

教育目的や、公的統計の二次的利用の際のテストデータなどの利用が可能。

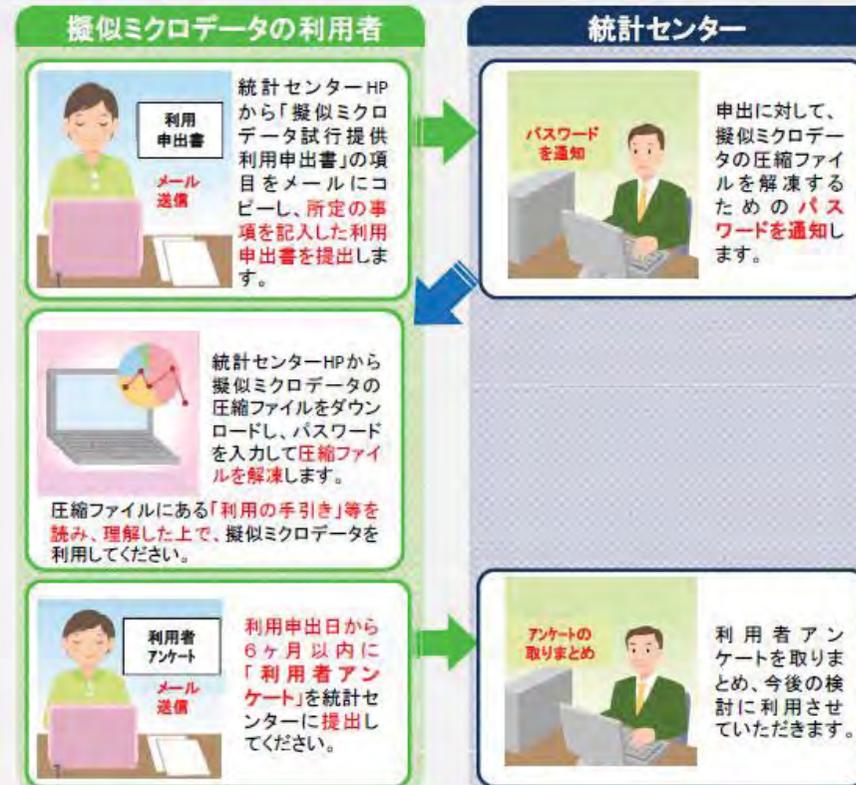
#### 【利用要件】

- ・申請者及び利用者以外の者に利用させないこと
- ・集計表から擬似的に作成したデータであるので、分析結果は実証研究の結果と見なすことができないことを理解すること
- ・利用者アンケートの提出

#### 統計センターが試行提供している擬似マイクロデータ(平成25年8月末現在)

- 平成16年全国消費実態調査に関する試行版の擬似マイクロデータ
  - \* 大規模データ (CSV形式のみ)  
約200項目:世帯情報、収入、実支出(消費支出、非消費支出)、実支出以外、繰越金  
約3万2千レコード
  - \* 簡易データ (CSV形式 及び Excel形式)  
25項目:世帯情報、消費支出追加  
約8千レコード

## (参考) 擬似マイクロデータの利用手続きの流れ

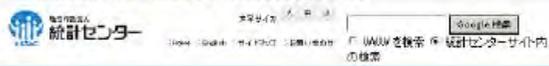


擬似マイクロデータの利用に関するFAQ: <http://www.nstac.go.jp/services/faq-gijimicro.html>

## (参考) 公的統計の二次的利用ホームページ

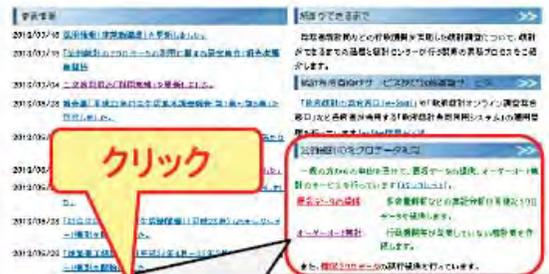
### ● (独) 統計センターホームページ

http://www.nstac.go.jp/



統計センターについて 業務案内 調査情報 情報公開 採用情報

統計センターは、国勢調査や国勢調査データなど、我が国の基本となる統計の作成(製表)を行っている独立行政法人です。



クリック

#### 公的統計のマイクロデータ利用

一般の方からの申出を受けて、匿名データの提供、オーダーメイド集計のサービスを行っています(インプレット)。

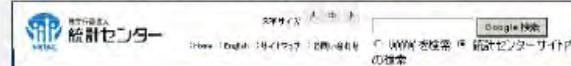
**匿名データの提供:** 多変量解析などの実証分析が可能なマイクロデータを提供します。

**オーダーメイド集計:** 行政機関等が公表していない統計表を作成します。

また、疑似マイクロデータの提供を行っています。

### ● 公的統計の二次的利用ホームページ

http://www.nstac.go.jp/services/archives.html



統計センターについて 業務案内 調査情報 情報公開 採用情報

統計センターの二次的利用サービスについて



# 統計センターにおける二次的利用基盤（概念図）



## 4 統計センターと学術研究機関等との連携について

### 学術研究機関等との連携協力の取組

統計センターでは、平成21年4月施行の統計法において創設された公的統計の二次的利用制度の充実と学術研究の発展を図っていくため、学術研究機関等と連携した取組を展開。公的統計の利用拡大に係る取組に賛同する法人と、連携協力協定を締結し、公的統計の二次的利用に関する研究・開発、普及・啓発を推進するほか、これらの法人に統計データアーカイブのサテライト機関の役割を担ってもらい、研究者等に向けた二次的利用のサービスの充実を図っている。

#### 法人の要件

1. 国立大学法人法に基づき設置された国立大学法人及び大学共同利用機関法人
2. 私立学校法に基づく学校法人により設置された私立大学
3. 独立行政法人通則法及び個別法の定めるところにより設立された独立行政法人
4. 法人税法別表第1に掲げる公共法人
5. 公益社団法人及び公益財団法人の認定等に関する法律により公益性の認定を受けた公益法人（特例民法法人を含む。）

#### 連携協力事項

- 公的統計に関するデータアーカイブの運営に関すること（**施設基準に適合**）
  - ・匿名データの提供
  - ・オンサイト利用環境の提供など
- 公的統計の二次的利用に関する研究・開発
- 公的統計の二次的利用に関する普及・啓発
- 人材交流
- その他協定の目的を達成するために必要な事項

#### 4 統計センターと学術研究機関等との連携について

### 連携協力協定を締結している大学等

サテライト機関名	匿名データの提供	オンライン利用環境の提供	二次的利用に係るURL
国立大学法人 一橋大学 経済研究所附属社会科学統計情報研究センター	○	○	<a href="http://rcisss.ier.hit-u.ac.jp/Japanese/micro/index.html">http://rcisss.ier.hit-u.ac.jp/Japanese/micro/index.html</a>
国立大学法人 神戸大学大学院 経済学研究科	○	—	<a href="http://www.econ.kobe-u.ac.jp/kuma/satellite/index.html">http://www.econ.kobe-u.ac.jp/kuma/satellite/index.html</a>
法政大学 日本統計研究所	○	—	<a href="http://www.hosei.ac.jp/toukei/micro/index.html">http://www.hosei.ac.jp/toukei/micro/index.html</a>
大学共同利用機関法人情報・システム研究機構 新領域融合研究センター(統計数理研究所)	○	○	<a href="http://www.rois.ac.jp/tric/tokumei/tokumei.html">http://www.rois.ac.jp/tric/tokumei/tokumei.html</a>

## 統計におけるオープンデータの高度化



- ▶ 政府統計の中核的機関である総務省統計局は、(独)統計センターと協力し、大量・多様な統計データの提供方法を次世代化し、データの高度利用を可能とする以下の取組を実施  
オープンデータ推進のトップランナーとして政府の取組を先導

1. API機能による統計データの高度利用環境の構築

2. 統計GIS機能の強化

3. オンデマンドによる統計作成機能・方策の研究

- ▶ これにより、官民における統計データ利活用の高度化を促進し、新たな付加価値を創造するサービスや革新的な事業の創出などを支援

## 1 API機能による統計データの高度利用環境の構築

—大量・多様な統計データの機械的処理—

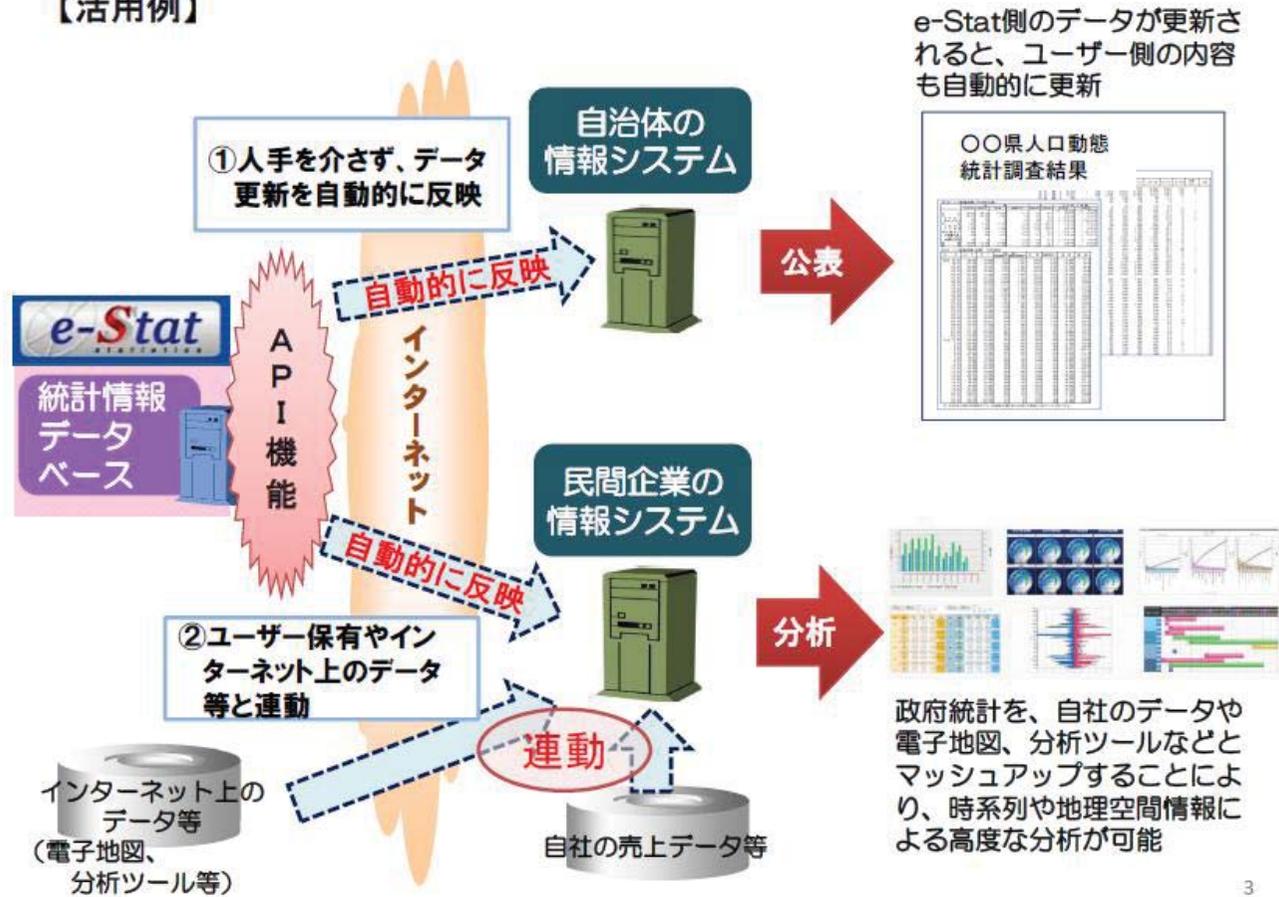
- ▶ 政府統計のポータルサイト『e-Stat』に、新たにAPI機能(Application Programming Interface)を付加するとともに、蓄積された統計データを機械判読可能な形式に変更することで、次のようなことが可能に

- ① 利用者の情報システムにe-Statのデータを自動的に反映
- ② ユーザー保有やインターネット上のデータ等と連動させた高度な統計データ分析



- ▶ 統計データ利用の高度化や効率化の環境が構築され、ビジネス活性化や新規事業の開発促進、行政サービス向上などに貢献
- ▶ 6月上旬からユーザー登録制で試行提供を開始予定  
(統計局所管の統計データによる試行提供によって機能やシステム負荷の検証を実施)

【活用例】



## 2 統計GIS機能の強化

- ▶ e-Statの統計GIS機能を強化し、ユーザー保有データの取り込み分析や任意に指定したエリアにおけるデータが利用可能になるような機能を開発中

- 【活用例】
- ①ユーザーの保有するデータを取り込んで分析する機能の追加
  - ②任意に指定したエリアにおける統計算出機能の追加

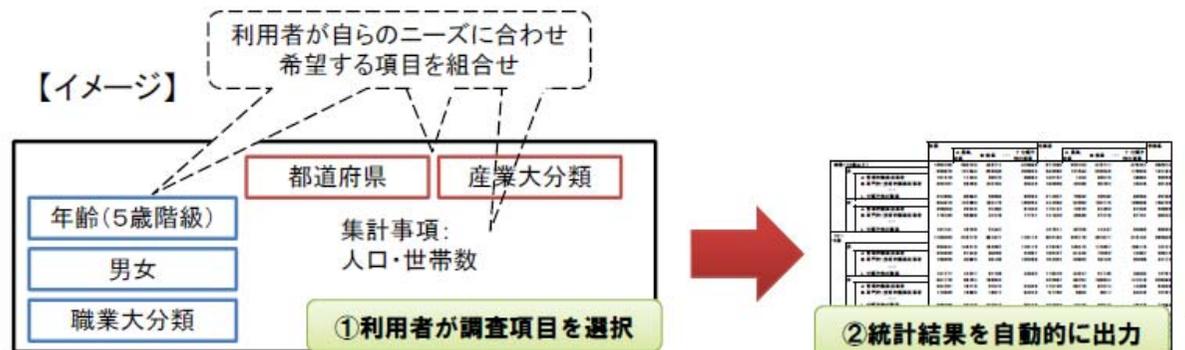


- ▶ 本年秋を目途にユーザー登録制で試行提供を開始予定  
(統計局所管の統計データによる試行提供によって機能やシステム負荷の検証を実施)

### 3 オンデマンドによる統計作成機能・方策の研究

- ▶ 公的機関や学術研究などの利用において、利用者が調査項目を選択するだけで統計結果を自動的に出力する、新しい形の統計提供サービスを研究中
- ▶ これにより、既存の結果表にない任意の多重クロス集計が出力可能になり、学術研究を始めとする多様なニーズに対応

※実用化に向けては運営・制度面及び秘匿処理における検討課題がある。



## (参考) 統計データのオープン化の現状



(平成25年5月時点)

- ▶ 統計結果は、原則インターネット上で公開
- ▶ 統計局所管統計の結果は、全て公開済み
  - データベースやスプレッドシート形式など、データとして利用可能な形式
  - 報告書には掲載しきれない結果データもインターネット上で公開

- ・統計結果はe-Statにスプレッドシート形式などで掲載することが原則(※)  
※統計調査等業務の業務・システム最適化計画(平成18年3月 各府省CIO連絡会議決定)
- ・一部府省では、PDFなどユーザーが自由に編集・加工できない形式での公開にとどまる統計もある

## 附属資料Ⅱ 海外における情報通信技術の活用に関する取組

### 1. アメリカセンサス局(United States Census Bureau)

・Centurion 概要



**U.S. Census Bureau**

**Centurion Internet Data Collection**

**Overview**

**John Studds**  
Information Technology (IT) Directorate

*June 20, 2013*



出所)US Census 提供資料

## Centurion Internet Data Collection – Overview

### Centurion - What is it?

The Centurion secure Internet data exchange system is a web-based framework for the design, delivery, and execution of surveys, censuses, and other data collection and data exchange efforts over the Internet in a secure manner.

The system allows for dynamic response applications, interactive web forms, fillable forms, file transfers/uploads, and data exchanges as mechanisms to effectively furnish the Census Bureau with higher quality data via the Internet.

These mechanisms have a full host of features to not only enhance the user experience, but to also reduce user burden and enhance data quality. The scalable, enterprise-level solution offers survey areas the ability to reach a large number of potential respondents on-line in a truly flexible and customizable manner to suit their business needs.

Features include:

- flexible appearance and functionality
- real-time edits and feedback messaging
- online help functionality - utilizing embedded screen timers that can automatically launch help functionality (be it static, readable text, audio help)
- file imports/exports,
- multi-language support
- dynamic content generation
- viewing & printing of provided data in form views
- dashboards
- review screens
- and other features.



United States  
**Census**  
Bureau

U.S. Department of Commerce  
Economics and Statistics Administration  
U.S. CENSUS BUREAU

# Centurion Internet Data Collection – Overview

## Flexible & Full-Featured – Improving Data Quality & Reducing Burden

The image displays a collage of various web-based data collection interfaces, illustrating the flexibility and full-featured nature of the Centurion system. The interfaces shown include:

- CDC Centers for Disease Control and Prevention:** A form for hospital admission and discharge information, including fields for patient name, date, and hospital details.
- National Science Foundation (NSF):** A survey titled "NATIONAL SURVEY OF COLLEGE GRADUATES" with a section for "Other Work-Related Experiences".
- 2012 Economic Census:** A form for "Form Review" with a list of steps to complete, such as "Step 1 - Employee Identification Number" and "Step 2 - Physical Location".
- United States Census Bureau:** The "American Community Survey" login page, featuring a barcode and instructions for users.
- State of Georgia:** The "2013 Annual Survey of Public Employment & Payroll" login page, including a "Welcome" message and login fields.

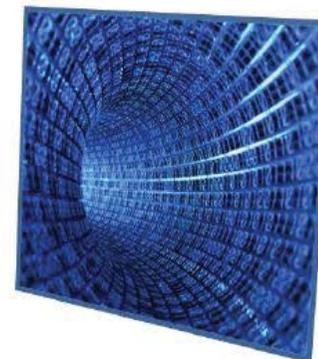
At the bottom left, the **United States Census Bureau** logo is displayed, along with the text: "U.S. Department of Commerce, Economics and Statistics Administration, U.S. CENSUS BUREAU".

## Centurion Internet Data Collection – Overview

### Benefits

Reduction of per survey/census development costs over previous Internet-based collection efforts by nearly one-third and the elimination of duplicate efforts by optimizing implementation

- Standardized, repeatable approach
  - Framework approach significantly reduces development time
    - Survey design & implementation
    - Database schemas
    - XML output format for collected data
    - Administration & Project Management
    - IT Security
  - Survey “Viewer” tool provides subject matter experts the ability to review Internet pages as they are being developed
  - Capability to use ‘paper-form’ design metadata and other survey content sources to automatically generate question screens, with validations and ‘skips’
- Further cost savings can be realized as the framework may be utilized by the Census Bureau’s telephone call centers for data collection during telephone follow-up activities, telephone questionnaire assistance, and CATI (11 surveys),
  - Dynamic response applications and interactive web forms implemented on Centurion for self-response over the Internet can also be used by the call centers with little or no modification,
- Standard Desktop and Mobile web browser clients (Smart phones, Blackberries, iPhones, etc) may be used to access the system. Centurion is also flexible enough that it can be ported to other devices such as laptops, tablets, etc.



United States  
**Census**  
Bureau

U.S. Department of Commerce  
Economics and Statistics Administration  
U.S. CENSUS BUREAU



# Centurion Internet Data Collection – Overview

## Metadata Translation & Rendering – Optimizing Implementation

For implementation purposes, we have developed a number of tools to make implementation more efficient, including a metadata translation and rendering engine allowing us to automatically generate screens by re-using metadata from various sources. Metadata about paper form creation, data elements, and content can all be extracted from repositories:

- Services Annual Survey -160 unique survey forms
- 2012 Economic Census more than 830 unique forms (each form consisting of more than 30 questions containing anywhere from 75 to 400 data items).

This approach makes the development process much more efficient and we continue to expand these capabilities out with new collections.

The image displays three overlapping screenshots related to data collection. On the left is a table with columns for 'FORM\_ID', 'FORM\_NAME', 'FORM\_NUMBER', 'REQUEST\_NUMBER', 'CATEGORY', 'FORM\_SECTION', 'RESPONSE\_EPISODE', 'RESPONSE\_LABEL', 'RESPONSE\_ATTRIBUTE\_GROUP', and 'RADIO\_GROUP'. The table contains multiple rows of data for '2011' forms. In the center is a screenshot of the '2011 Annual Services Report' form, which includes a 'Welcome to the Annual Services' message, a 'Log in by entering the appropriate information' section with fields for 'Username:' and 'Password:', and a 'Please note: sessions will expire (required)' warning. On the right is a screenshot of the '2012 Economic Census' login page, featuring the 'United States Census' logo, a 'Welcome to the 2012 Economic Census' message, a 'Log in' section with 'User ID:' and 'Password: (case sensitive)' fields, and a 'Log in' button. The background of the entire image is a collage of various scenes, including people and buildings.

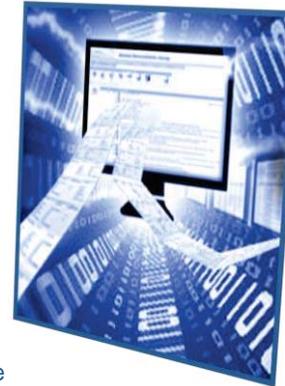
United States  
**Census**  
Bureau

U.S. Department of Commerce  
Economics and Statistics Administration  
U.S. CENSUS BUREAU

## Centurion Internet Data Collection

### Benefits (continued)

- Assist survey areas in enhancing or modernizing key business processes and meeting or exceeding their goals, these data collection efforts may now utilize “custom” features to facilitate their unique on-line data exchange requirements, as well as, the many standardized components already developed,
- Time and cost-effective strategies can easily be integrated into the flexible Centurion framework to help address functional gaps and/or realize new efficiencies
  - Streamlined data processing and post data collection activities
  - Modified mailing and contact strategies in development based on high Internet response rates
    - 6 data collections receive more than 70% of their total responses via the Internet
    - 10 additional data collections receive more than 50% of their total responses via the Internet
- Collection of data is more cost effective and has a higher degree of accuracy when compared to equivalent traditional collection methods
  - Real-time ‘item specific’ edits, validations, and messaging for higher quality data receipts upon submission (e.g., ACS Study – Date of Birth screen)
  - Content and question specific ‘Help’ available to respondents (multiple variations),
  - Dynamic question content generated based on previous responses or prior period data
  - Response review and correction functionality prior to submission,



United States™  
**Census**  
Bureau

U.S. Department of Commerce  
Economics and Statistics Administration  
U.S. CENSUS BUREAU



## Centurion Internet Data Collection – Overview

### Security

- The Office of Information Security is actively engaged in the Centurion framework and the implementation of ALL features, modules, and data collections
  - Centurion is fully Certified & Accredited with NO outstanding POAMs and is part of the continuous monitoring effort
- Third-party performance, penetration, and security testing
- The Centurion software architecture uses the industry-standard Model-View-Controller (MVC) design, and emphasizes a high degree of modularity and configurability, but ensures security controls are applied consistently throughout the data collections.
- All communications between the data provider and Centurion take place over a secure HTTPS connection.
- Segregation of systems and databases
- All data saved, stored, and retained by the system is stored in an encrypted form within the database tier.
- NIST security controls are in place as identified in NIST SP 800-53 and in the Census Bureau's organizational baseline security requirements.



United States  
**Census**  
Bureau

U.S. Department of Commerce  
Economics and Statistics Administration  
U.S. CENSUS BUREAU



## Centurion Internet Data Collection – Overview

### Security - Auditing

- A full host of security auditing, logging, and scanning are employed by the system and supporting systems at all times. Logs are automatically provided to the enterprise-level 'Sensage System' and Centurion Control Center for monitoring and review.
- Centurion has in place existing user-action monitoring and auditing.
- Data consistency monitoring is employed once data providers have submitted data in an attempt to detect any irregularities in data that may indicate suspicious or anomalous activities.
- Enterprise-level HP Application Management is in place to monitor the health of the systems that support Centurion and provide alerts as needed.
- Data from the Centurion system are replicated in real-time and are also backed up nightly to FIPS 140-2 encrypted media. These media are then stored in two distinct locations.



United States  
**Census**  
Bureau

U.S. Department of Commerce  
Economics and Statistics Administration  
U.S. CENSUS BUREAU

# Centurion Internet Data Collection – Overview

## Security By-products

Centurion additionally supports convenient access to tracking respondent progress, data, and survey process data, empowering a high level of transparency into the ongoing status of Internet-based data collection efforts

- Paradata (available as a by-product of the system's security features)
- Respondent information & status
- Respondent data views,

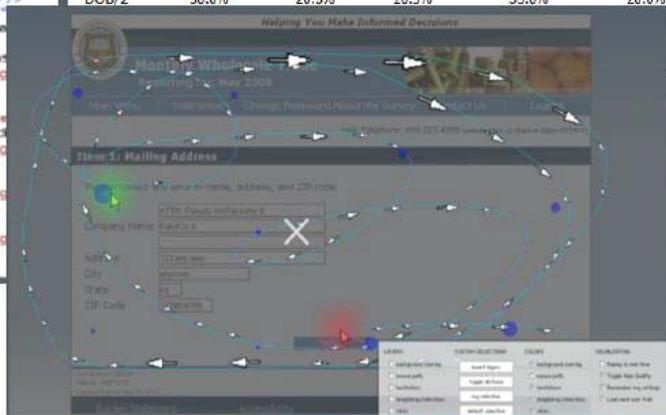
```

<dataset type="paradata" service="GUS">
  <respondent id="###Unique Respondent id###">
    <event type="failed_login" time="1323380993" page="login">
    <event type="login" time="1323381039" page="login">
      <environment width="1900" userAgent="Mozilla/4.0 (com
    </event>
    <event type="entry" time="1323381039" page="intro"/>
    <event type="exit" time="1323381045" page="intro"/>
    <event type="hyperlink" time="1323381045" page="intro">
      <link url="https://respond.census.gov/gus/cea
    </event>
    <event type="entry" time="1323381046" page="ceas
    <event type="field_change" time="1323381058" pag
      <field value="1" name="ceased_effect"/>
    </event>
    <event type="next_action" time="1323381065" pag
    <event type="entry" time="1323381066" page="back
    <event type="field_change" time="1323381091" pag
      <field value="1" name="depend"/>
    </event>
    <event type="field_change" time="1323381109" pag
      <field value="2" name="website"/>
    </event>
    <event type="field_change" time="1323381117" pag
      <field value="3" name="fisc_year_1"/>
    </event>
  </respondent id="###Unique Respondent id###">
</dataset>

```

Errors due to blanks

Screen	% Completely Blank	% DOB Blank (age provided)	% Age Blank (invalid DOB)	% Completely Blank, Corrected	% Completely Blank, Age (DOB still blank)	% DOB Blank, Corrected	% Age Blank, Corrected
DOB	27.2%	25.4%	8.7%	54.5%	15.8%	64.2%	85.6%
DOB/1	21.7%	29.2%	5.0%	60.0%	11.5%	82.4%	84.2%
DOB/2	30.0%	20.3%	10.9%	53.6%	20.0%	49.5%	86.3%
						45.3%	90.0%
						34.3%	88.9%
						18.2%	66.7%



U.S. Department of Commerce  
Economics and Statistics Administration  
U.S. CENSUS BUREAU



# Centurion Internet Data Collection - Overview

## Monitoring

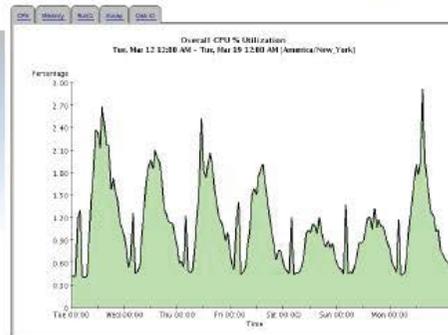
Data consistency monitoring is employed once data providers have submitted data in an attempt to detect any irregularities in data that may indicate suspicious or anomalous activities.

### Bowie Computer Center – Enterprise Operations Center

- provides 24x7x365 monitoring services of systems and devices on the Bureau of the Census (BOC) network.
- provides BOC customers with server performance statistics such as CPU, memory and other key utilization metrics
- provides newly enhanced Enterprise-level HP Application Management to monitor the health of the systems that support Centurion and provide 24x7x365 alerts, notifications, and escalations as needed.

#### IT System Resource Past 7 Days

Click on Column and Open Graph of the resource, CPU, Memory, Run Queue, Swap, and Disk I/O utilization for a given day and time. Detailed data will be available.



Alert	Subject	Severity	Time
07140568-801c-884c-2068-207041171549	HP Application Management Alert: Critical Availability Alert: Severity: Critical	Critical	2013-03-19 15:42:59.04
07140568-801c-884c-2068-207041171549	HP Application Management Alert: Critical Availability Alert: Severity: Critical	Critical	2013-03-19 15:42:59.04
07140568-801c-884c-2068-207041171549	HP Application Management Alert: Critical Availability Alert: Severity: Critical	Critical	2013-03-19 15:42:59.04
07140568-801c-884c-2068-207041171549	HP Application Management Alert: Critical Availability Alert: Severity: Critical	Critical	2013-03-19 15:42:59.04
07140568-801c-884c-2068-207041171549	HP Application Management Alert: Critical Availability Alert: Severity: Critical	Critical	2013-03-19 15:42:59.04
07140568-801c-884c-2068-207041171549	HP Application Management Alert: Critical Availability Alert: Severity: Critical	Critical	2013-03-19 15:42:59.04
07140568-801c-884c-2068-207041171549	HP Application Management Alert: Critical Availability Alert: Severity: Critical	Critical	2013-03-19 15:42:59.04
07140568-801c-884c-2068-207041171549	HP Application Management Alert: Critical Availability Alert: Severity: Critical	Critical	2013-03-19 15:42:59.04
07140568-801c-884c-2068-207041171549	HP Application Management Alert: Critical Availability Alert: Severity: Critical	Critical	2013-03-19 15:42:59.04



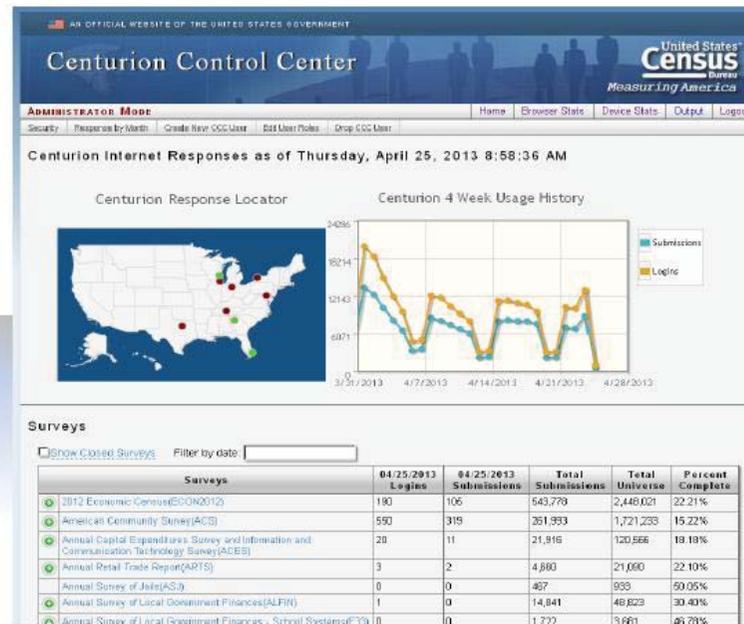
U.S. Department of Commerce  
Economics and Statistics Administration  
U.S. CENSUS BUREAU

## Centurion Internet Data Collection - Overview

### Centurion Control Center

Centurion employs a number of 'real-time' monitoring capabilities:

Auditing, logging, and user-interaction data are provided in real-time to the Centurion Control Center and the enterprise Sensage system for analysis, action, and review.



U.S. Department of Commerce  
Economics and Statistics Administration  
U.S. CENSUS BUREAU

## Centurion Internet Data Collection – Overview

### Accessible

Standard Desktop and Mobile web browser clients (Smart phones, Blackberries, iPhones, etc) may be used to access the system. Centurion is also flexible enough that it can be ported to other standalone-devices such as laptops and tablets.

**Centurion User Device Info as of Mar 20, 2013 12:00**

Rank	Device	Logins	Percent
1	Other	1,893,578	98.01%
2	Flad	27,798	1.41%
3	Android	6,499	0.33%
4	Phone	4,271	0.22%
5	YVAF capable phone	44	0.00%
6	BlackBerry	42	0.00%
7	Sony PlayStation 3	10	0.00%
8	Nokia Lumia 900	8	0.00%
9	Nokia Lumia 710	8	0.00%
10	HTC Flyer 4G	7	0.00%
11	SAMSUNG SCH-B27	6	0.00%
12	Amazon Kindle	6	0.00%
13	HTC T3295	5	0.00%
14	HTC HD7	4	0.00%
15	webOS	4	0.00%
16	HTC T9788	3	0.00%
17	SAMSUNG SCH-B97	2	0.00%
18	HTC PC8100	2	0.00%
19	Nintendo 3DS	1	0.00%
20	HTC P96100	1	0.00%
21	SAMSUNG SCH-B17	1	0.00%

**Centurion User Browser Info as of Mar 20, 2013 12:00**

Browser	Version	Logins	Percent
MSE	All Versions	1,397,583	67.83%
Firefox	All Versions	266,667	13.48%
Chrome	All Versions	231,463	11.74%
Safari	All Versions	116,819	5.69%
ACL Browser	All Versions	14,207	0.72%
Mobile Safari	All Versions	3,725	0.19%
RealPlayer	All Versions	2,215	0.11%
Opera	All Versions	809	0.04%
Netscape	All Versions	68	0.00%
NetFront	All Versions	33	0.00%
BlackBerry	All Versions	27	0.00%
IceWeasel	All Versions	24	0.00%
WakaTV	All Versions	1	0.00%
Nintendo 3DS	All Versions	1	0.00%
ELink	All Versions	1	0.00%

**Top 10 Centurion Browser Versions**

Rank	Browser	Version	Logins	Percent
1	MSE	8	603,960	30.89%
2	MSE	9	548,611	27.82%
3	MSE	7	195,462	7.88%
4	Chrome	24	78,931	4.00%
5	Firefox	18	72,415	3.67%



U.S. Department of Commerce  
Economics and Statistics Administration  
U.S. CENSUS BUREAU





## Centurion Internet Data Collection - Overview

### Moving Forward

Continued introduction of enhanced reporting and respondent burden-reducing tools & techniques,

Continued introduction of enhanced data collection implementation tools,

Continued introduction of new monitoring and provisioning tools,

Continued expansion of the metadata repository and shared services

Implementation of additional capabilities to further empower business-line, subject matter experts to realize new cost and time saving efficiencies,

Expansion of data collection operations:

- Data collections currently implemented = 64.
- More than 1160 'paper-based' questionnaires represented.
- Number respondents offered the option to report via the system is now over 10.9 million.
- More than 3.2 million finalized responses have been received.
- International data collection efforts
- Offering Expanded Collection Services to Others

**KIMBERLEY PROCESS**  
ROUGH DIAMOND STATISTICS

**PUBLIC STATISTICS AREA**

In accordance with the terms of the Administrative Decision on the Public Release of Aggregated Statistical Data, the following table displays statistical data of Kimberley Process Participants for trade and production of rough diamonds, by value and volume, and Kimberley Process Certificate counts.

These data are considered public information; however, the explanatory notes provided below should be reviewed whenever these data are to be used.

Statistical data will be made available approximately six months after the end of each reference year. Requests for previously published data may be made at that time as warranted.

- Administrative Decision on Publication of Aggregated Statistical Data
- Explanatory Notes for Public Release of Kimberley Process Certificate Statistics Data

Report Type	Report Year
Annual Summary Tables	2010 2009 2008 2007 2006 2005 2004
Annual Summary Charts	2010 2009 2008 2007 2006 2005 2004
State/Annual Production Tables	2010 2009 2008 2007 2006 2005 2004
Quarterly Export Tables	2010 2009 2008 2007 2006 2005 2004
Quarterly Import Tables	2010 2009 2008 2007 2006 2005 2004
Quarterly Certificate Count Tables	2010 2009 2008 2007 2006 2005 2004

United States<sup>™</sup>  
**Census**  
Bureau

U.S. Department of Commerce  
Economics and Statistics Administration  
U.S. CENSUS BUREAU

2. アメリカ労働統計局(United States Bureau of Labor Statistics)

1)CES 調査票例(製造業、事業所 2 箇所以上の企業向け)

► Our records show the following information for your firm:

<b>Contact:</b>	<b>Report Number:</b>
<b>Tel: Ext:</b>	<b>Location:</b>
<b>Fax:</b>	<b>UI Number:</b>
	<b>Industry Code:</b>
	<b>Email:</b>

► Definitions for the Questions on the Next Page:

**Column 1 EMPLOYEE COUNT– ALL WORKERS**

Total number of persons in this pay group who worked or received pay for any part of the pay period that includes the 12<sup>th</sup> of the month.

**Include:**

- Executives and their staff
- Full-time and part-time workers
- Salaried officials of corporations
- Trainees
- Workers on active duty, if receiving pay from employer
- Workers on paid sick leave
- Workers on paid vacation
- Workers on other paid leave

**Exclude:**

- Outside contractors and their employees
- Pensioners
- Proprietors, owners, or partners of unincorporated firms
- Workers on active duty, if not receiving pay from employer
- Workers on leave without pay for entire pay period
- Workers on strike for entire pay period
- Unpaid family members

**EMPLOYEE COUNT– PRODUCTION WORKERS**

Number of "All Workers" defined above who are Production Workers. Production Workers include working supervisors or group leaders who may be "in charge" of some employees, but whose supervisory functions are only incidental to their regular work.

**Include individuals working in:**

- Assembling
- Fabricating
- Janitorial activities
- Maintenance or repair
- Materials handling
- Processing
- Product development
- Recordkeeping related to production
- Shipping or receiving
- Storage or warehousing
- Trucking

**Exclude individuals working in:**

- Accounting or finance
- Advertising
- Cafeterias
- Collection and credit
- Executive, professional, or technical positions
- Force account construction
- Legal
- Medical
- Personnel
- Product installation or servicing
- Purchasing
- Recordkeeping not related to production
- Sales and delivery

**Column 2 WOMEN EMPLOYEE COUNT**

Number of "All Workers" defined above who are women.

**Column 3 PAYROLL, EXCLUDING COMMISSIONS**

Total gross pay earned during the entire pay period. Report separately for All Workers and for Production Workers.

**PAYROLL, EXCLUDING COMMISSIONS (CONTINUED)**

Report pay before employee deductions for:

- Taxes
- FICA (Social Security)
- Health insurance
- Pay deferral plans such as 401K
- Bonds
- Pensions
- Unemployment insurance
- Union dues

**Include:**

- Wages and salaries
- Paid holidays, vacation, sick leave, and other paid leave
- Incentive pay
- Bonuses paid each pay period
- Overtime pay
- Severance, if paid over multiple pay periods

**Exclude:**

- Commissions
- Annual pay for unused leave
- Awards or bonuses not paid each pay period
- Employer contributions to pay such as 401K
- Pay advances, such as vacation pay advances
- Payments "in kind"
- Retroactive or back pay
- Severance, if provided as one payment
- Travel or work-related reimbursements

**Column 4 COMMISSIONS**

Report separately for "All Workers" and for "Production Workers".

- Report for the most recent complete period for which commissions are available, which might be different from the pay period that includes the 12th. Enter 0 if none paid for the period or pay group.
- Exclude base pay, drawing accounts, or basic guarantees.

**Column 5 HOURS, INCLUDING OVERTIME**

Total number of hours for which employees received pay during the entire pay period. Report separately for All Workers and Production Workers.

- Include overtime; stand-by or reporting time; and hours not worked, but for which workers received pay (holidays, vacations, sick leave, etc.).
- Report hours for salaried and commission-only employees based on their standard work week.
- Do not convert overtime or other premium hours to straight-time equivalent hours.

**Column 6 OVERTIME HOURS**

Total number of hours for which employees received overtime premiums because they worked more than their regularly scheduled hours. Report separately for All Workers and Production Workers.

- Include Saturday, Sunday, 6<sup>th</sup> day, 7<sup>th</sup> day, and holiday hours.
- Exclude shift differential, hazard, incentive, or similar premiums.

Report Number:

Name of Firm:

Each month report your payroll information for the pay period that includes the 12<sup>th</sup> of the month. For questions refer to page 2 for the **Column** definitions or.

Employees receive pay:		Column 1	Column 2	Column 3	Column 4	Column 5	Column 6
Pay Group 1: Commissions Pay Group 1		EMPLOYEE COUNT	WOMEN EMPLOYEE COUNT	PAYROLL, EXCLUDING COMMISSIONS (Whole dollars)	COMMISSIONS, PAID AT LEAST ONCE A MONTH (Whole dollars)	HOURS, INCLUDING OVERTIME (Whole hours)	OVERTIME HOURS (Whole hours)
Pay period that includes January 12 <sup>th</sup> 2014	All Workers			\$	\$		
	Production Workers		N/A. Data not collected.	\$	\$		
February 12 <sup>th</sup> 2014	All Workers			\$	\$		
	Production Workers		N/A. Data not collected.	\$	\$		
March 12 <sup>th</sup> 2014	All Workers			\$	\$		
	Production Workers		N/A. Data not collected.	\$	\$		
April 12 <sup>th</sup> 2014	All Workers			\$	\$		
	Production Workers		N/A. Data not collected.	\$	\$		
May 12 <sup>th</sup> 2014	All Workers			\$	\$		
	Production Workers		N/A. Data not collected.	\$	\$		
June 12 <sup>th</sup> 2014	All Workers			\$	\$		
	Production Workers		N/A. Data not collected.	\$	\$		
July 12 <sup>th</sup> 2014	All Workers			\$	\$		
	Production Workers		N/A. Data not collected.	\$	\$		
August 12 <sup>th</sup> 2014	All Workers			\$	\$		
	Production Workers		N/A. Data not collected.	\$	\$		
September 12 <sup>th</sup> 2014	All Workers			\$	\$		
	Production Workers		N/A. Data not collected.	\$	\$		
October 12 <sup>th</sup> 2014	All Workers			\$	\$		
	Production Workers		N/A. Data not collected.	\$	\$		
November 12 <sup>th</sup> 2014	All Workers			\$	\$		
	Production Workers		N/A. Data not collected.	\$	\$		
December 12 <sup>th</sup> 2014	All Workers			\$	\$		
	Production Workers		N/A. Data not collected.	\$	\$		

2)IDCF(Internet Data Collection Facility)概要

# Internet Data Collection Facility (IDCF)

---

12/9/2013



[www.bls.gov](http://www.bls.gov)

出所)BLS 提供資料

# Collection Instruments

---

- Web form data entry with past data
  - ▶ User login account required
- Web form data entry without past data
  - ▶ User login account not required
- File transfer of bulk data sets
- E-mail routing
- Adobe fillable forms

# Surveys using IDCF

---

- Current Employment Survey (CES)
- Multiple Worksite Report (MWR)
- Job Openings and Labor Turnover (JOLTS)
- Annual Refiling Survey (ARS)
- Occupational Employment Statistics (OES)
- International Prices (IPP)
- Producer Prices (PPI)
- Occupational Injuries and Illnesses (SOII)
- National Compensation Survey (NCS)



# Surveys using IDCF

- Current Employment Survey (CES)
- Multiple Worksite Report (MWR)
- Job Openings and Labor Turnover (JOLTS)
- Annual Refiling Survey (ARS)
- Occupational Employment Statistics (OES)
- International Prices (IPP)
- Producer Prices (PPI)
- Occupational Injuries and Illnesses (SOII)
- National Compensation Survey (NCS)



# Benefits

---

- Real-time data edits as the respondent completes the survey reduces data errors and follow-up time/cost/burden
- Data enter internal processing databases faster and with fewer errors
- Lowest cost data collection method other than Touchtone Data Entry (TDE)



# Benefits

---

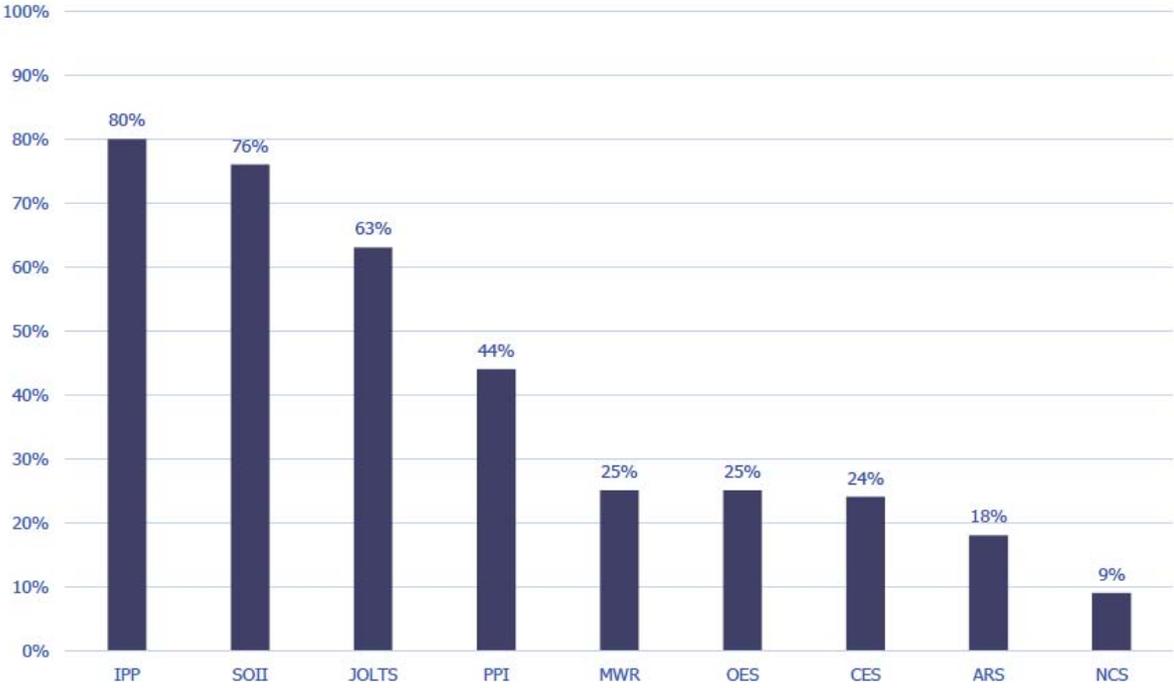
- New survey instruments can be deployed quickly and existing instruments can be easily modified
- Automated e-mail reminders to complete the survey that can direct respondents to the on-line collection instrument

# Challenges

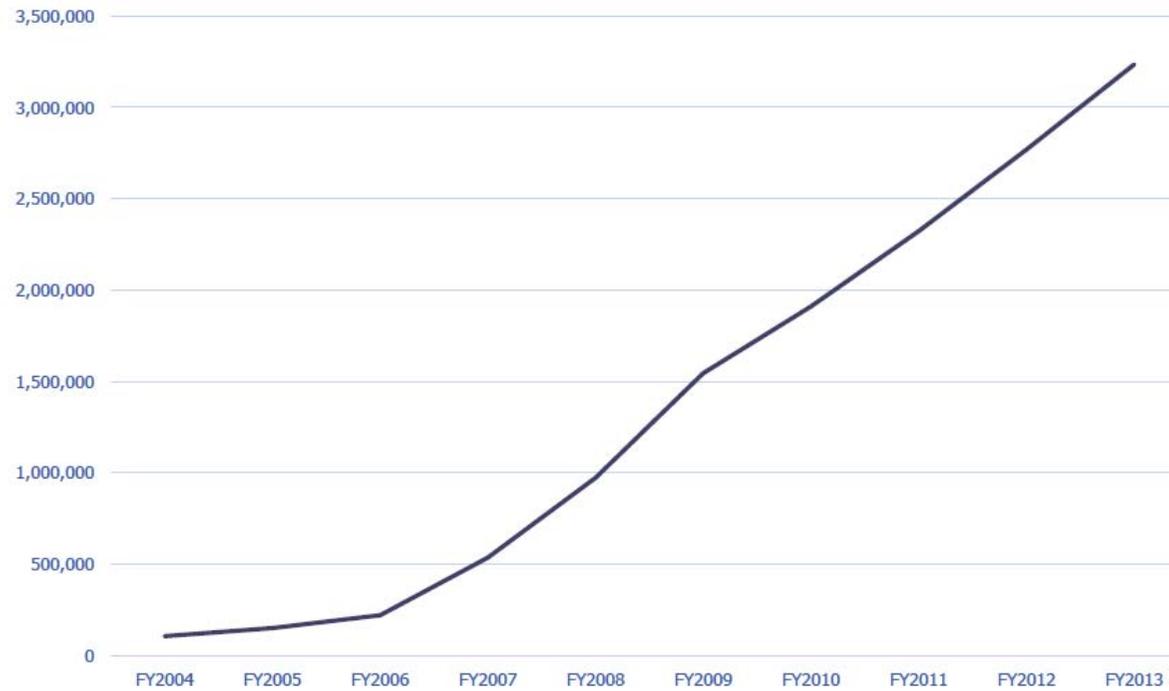
---

- Collection instruments have to be tested and verified to work with different web browsers and the many different versions of browsers used by survey respondents
- Respondents have varying Internet security controls within their systems
- Ensuring the high availability and performance of the websites

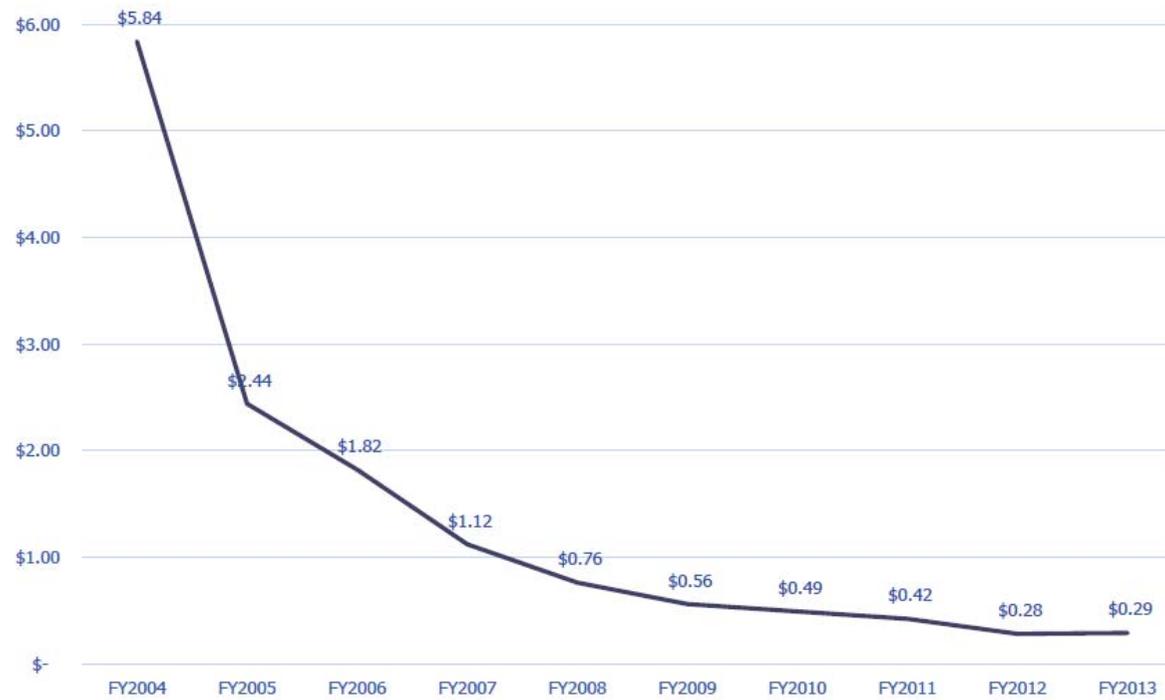
# FY2014 Percentage of Program Data via IDCF



# Units of Data Collected



# Cost per Unit of Data



# Units of Data Collected per HelpDesk Call



# Contact Information

---

**Rick Kryger**

Director of Technology and Computing  
Services

[www.bls.gov](http://www.bls.gov)

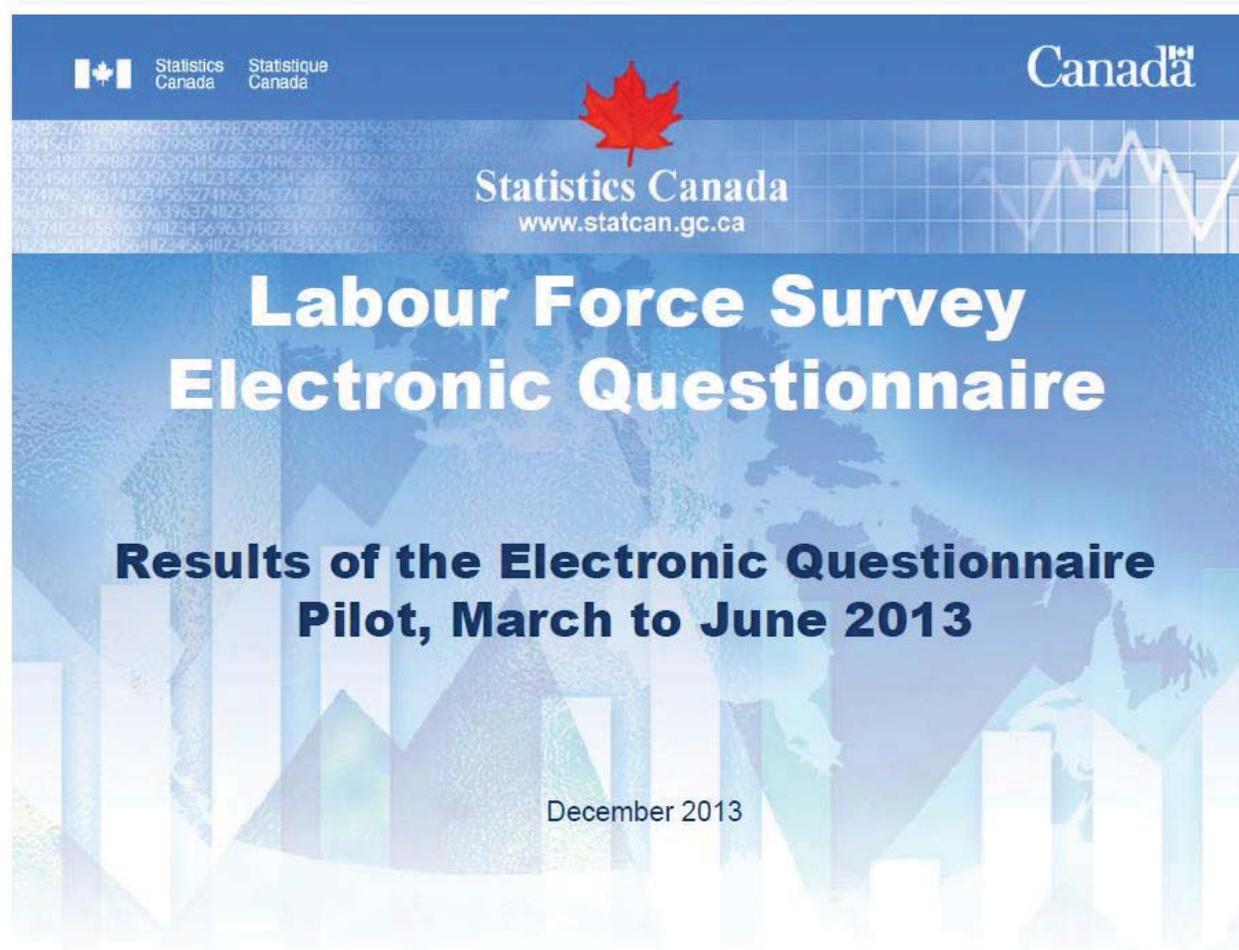
kryger.rick@bls.gov



[www.bls.gov](http://www.bls.gov)

### 3. カナダ統計局(Statistics Canada)

#### 1) LFS におけるインターネット回答の検証について



出所) カナダ統計局提供資料

## Outline

- Overview
- Goals
- Results and analysis
- Next steps



## LFS Overview

- Chosen dwellings are in sample for 6 consecutive months
- The collection window is 10 days
- The first month of collection is called a “birth”
- “Births” are collected using telephone or personal interviews
- After the first interview, most responses come from telephone interviews
- The Electronic Questionnaire (EQ) will be more convenient for respondents, and require less interviewers

## EQ Collection Strategy

- Typical birth collection
  - Telephone/Personal interview
  - At the end of interview, if respondent is eligible, ask if willing to answer via Internet for the following month
- Following months
  - EQ
    - 3 or 4 day window
    - Cases are converted to Telephone after EQ closes
  - Other cases are completed based on last month mode and try to convert to Internet at end of interview

## The Pilot Design

- Target: 6,400 households
  - Half introduced in March 2013 (3,608 cases)
    - Cases in sample for 4 months
  - Half introduced in April 2013 (3,601 cases)
    - Cases in sample for 3 months
    - collection ended in June
- Sample in New Brunswick, Québec, Ontario and Alberta
- Mostly urban except rural part of Ontario
  - Cases for the pilot were selected from clusters already in use in production

## EQ Eligibility

- Final outcome code of 70 (fully completed)
- Only one economic family in the dwelling
- The dwelling has a postal code that is 6 characters long, and a valid telephone number
- A current household member was identified as the household contact for the following month
- The interviewer completed the entire survey from beginning to end with the household contact for the following month
- There are no household members with fictitious names

## Goals of the LFS EQ pilot

- Operational feasibility
  - Length of EQ response window
  - Impact of email invitation and reminders on response rate
  - Response patterns (within EQ, and changes to telephone and personal collection requirements)
  - Measure collection costs (including helpdesk, coding, extra calls, and follow-up)
  - Partial responses, and proxy responses
  - Respondent feedback
- Application development
  - Processing system – edits, imputation, system time
  - EQ offer
  - Questionnaire design
- Impact on data quality
  - Includes short path, gross flows, non-response, and variable differences
  - Estimates & mode effect

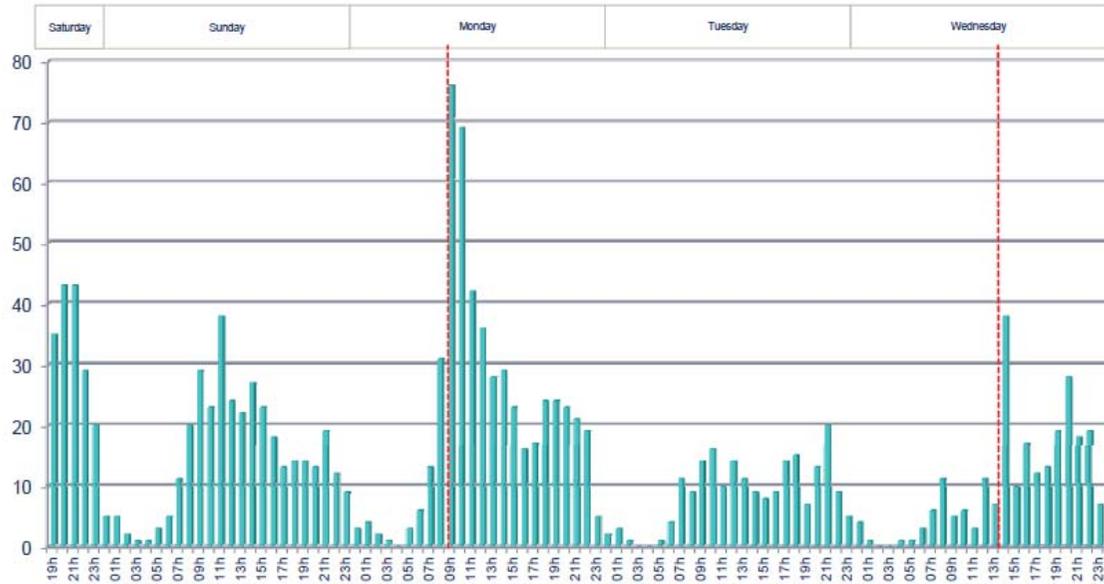
## LFS EQ eligibility, acceptance, and completion

Month	Rotation	Respondents (a)	EQ Eligibility (b)	EQ Acceptance (becomes the EQ sample) (c)	EQ Eligibility (b/a)	EQ Acceptance (c/a)
March	3	2,803	2,036	1,187	73%	42%
April	3 & 4	5,651	4,091	2,474	72%	44%
May	3 & 4	5,729	4,154	2,484	73%	43%

	Rot.	Total sample (d)	EQ sample (e)	Total respondents (f)	EQ respondents (g)	% of cases from EQ sample that were completed in EQ (g/e)	% of EQ sample cases that were completed, using any collection method	EQ Completion (of final respondents) (g/f)	EQ Completion (of total sample) (g/d)
April	3	7,114	1,187	5,651	527	44%	94%	9%	7%
May	3 & 4	6,996	2,474	5,729	1,228	50%	94%	21%	18%
June	3 & 4	6,985	2,484	4,831	1,464	59%	74%	30%	21%



# June submit times



### June submit times





### Collection “hours” for LFS-EQ and production cases

	Type	# of dwellings	Response rate	EQ respondents	CATI interviewers			
					Time (machine hours)	Attempts	Time / dwelling	Attempts / dwelling
April	PILOT	3,513	87%	527	884.8	13,952	0.31	3.97
	PROD	3,533	88%		976.8	15,366	0.34	4.35
May	PILOT	6,995	89%	1,228	1,487.7	27,693	0.21	3.96
	PROD	7,146	88%		1,816.7	34,137	0.25	4.78

92 hours difference in April (one rotation)  
329 hours difference in May (two rotations)

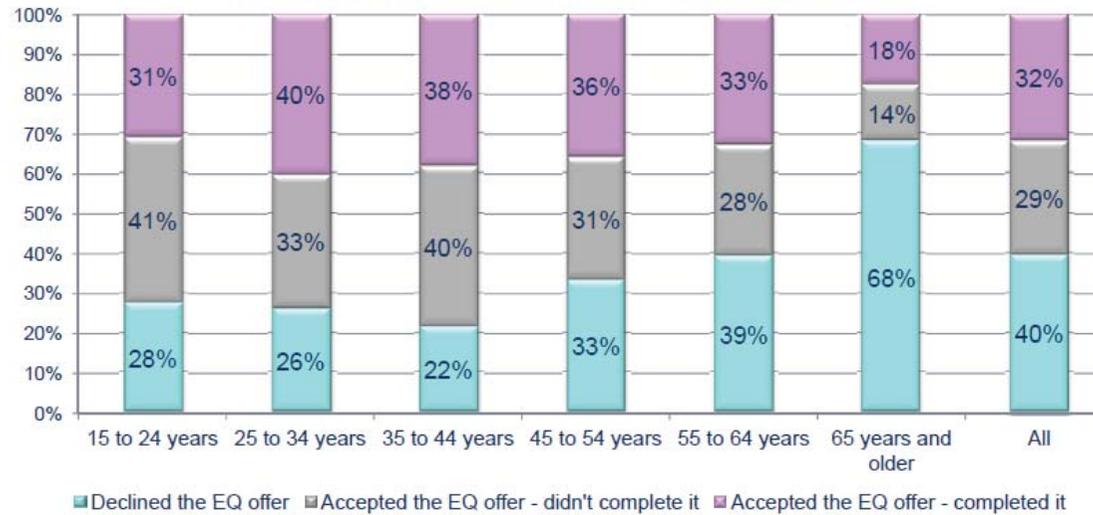


## Analysis and validation

- Characteristics of EQ respondents
- Estimates
- Distribution of industry and occupation
- Qualitative feedback
- Response rate
- Changes in household size
- Edit and imputations
- Short paths
- Changes in labour force status
- Question distributions
- Final variable distributions
- Final variable numerical averages
- Unemployment estimates

## Distribution of May EQ respondents, by age

Distribution of selected respondents by age group according to whether or not they accepted to respond to the EQ





## Estimates

- Collection strategy does not change the estimates
  - Not statistically different between production and pilot
- Examined 17 estimates in 4 provinces
- Subgroups:
  - Both sexes, 15+
  - Youths, 15-24
  - Men, 25+
  - Women, 25+
  - Core-age, 25-54

## Qualitative feedback

- From technical support, respondent comments, and interviewer debriefings
- The comments were related to:
  - Login process & passwords
  - Short collection period
  - Survey-specific details (spelling, variable questions)
  - EQ offer wording / training



## Analysis and validation

- The following items were also checked:
  - Response rate
  - Changes in household size
  - Edit and imputations
  - Short paths
  - Changes in labour force status
  - Question distributions
  - Final variable distributions
  - Final variable numerical averages
  - Unemployment estimates (still calculating)



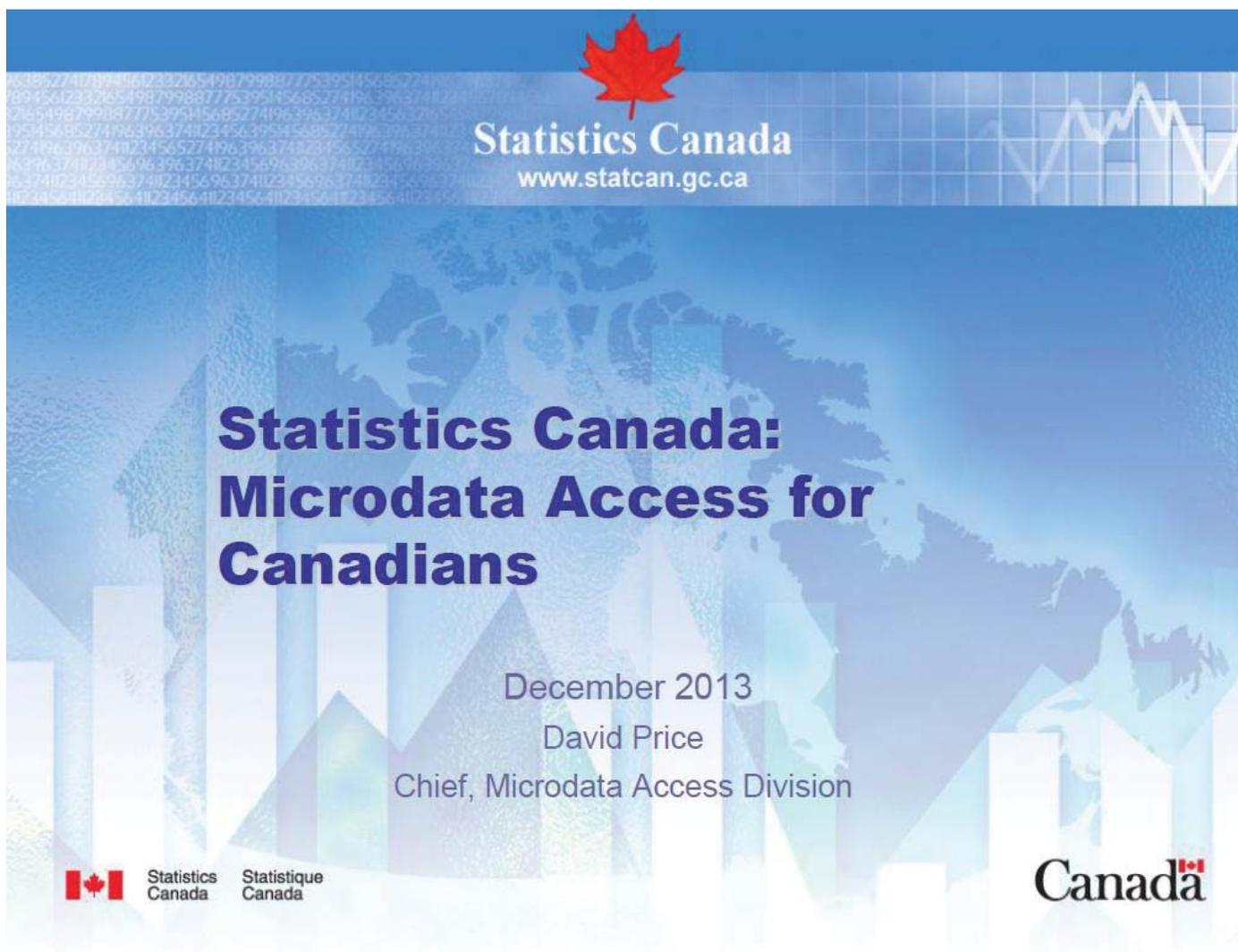
## Next Steps

- Mode effect results & final analysis
- Simplified EQ offer
- Technical challenges
- Roll-out strategy
- Full production in late 2014
- Monitor mode effect
- Long term: new integrated collection application with survey redesign

## Contact

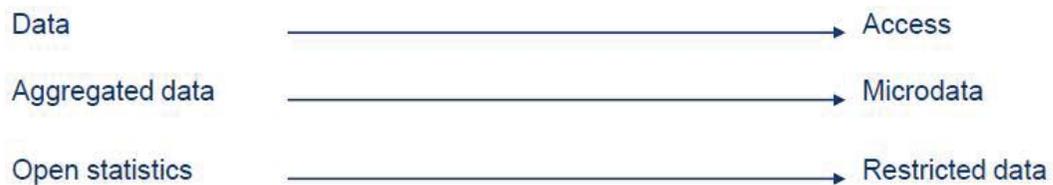
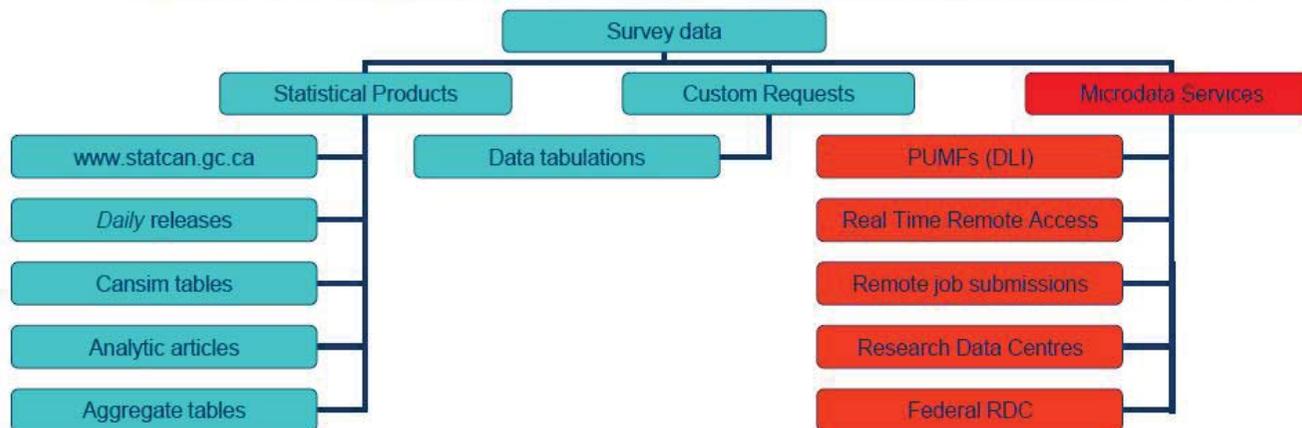
- Nicole MacDonald
  - [Nicole.MacDonald@statcan.gc.ca](mailto:Nicole.MacDonald@statcan.gc.ca)
- Marc Lachance
  - [Marc.Lachance@statcan.gc.ca](mailto:Marc.Lachance@statcan.gc.ca)

## 2) マイクロデータのアクセスについて



出所) カナダ統計局提供資料

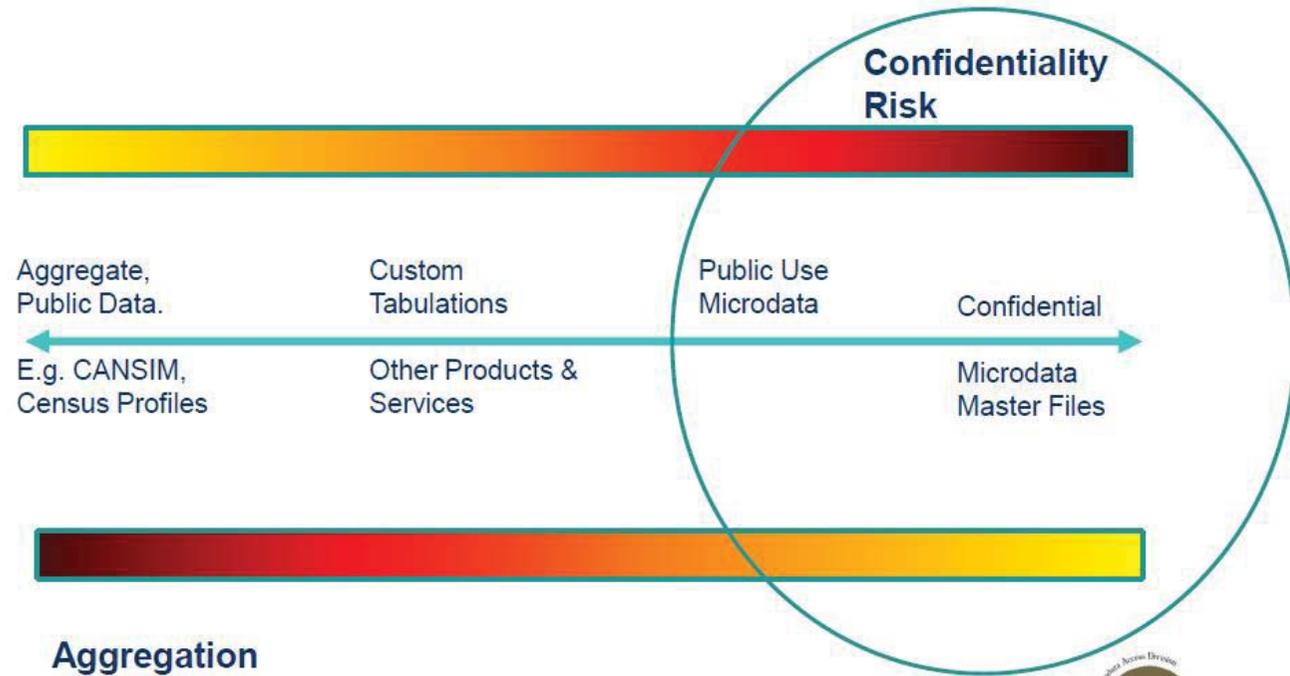
# Dealing with Data



Statistics Canada • Statistique Canada



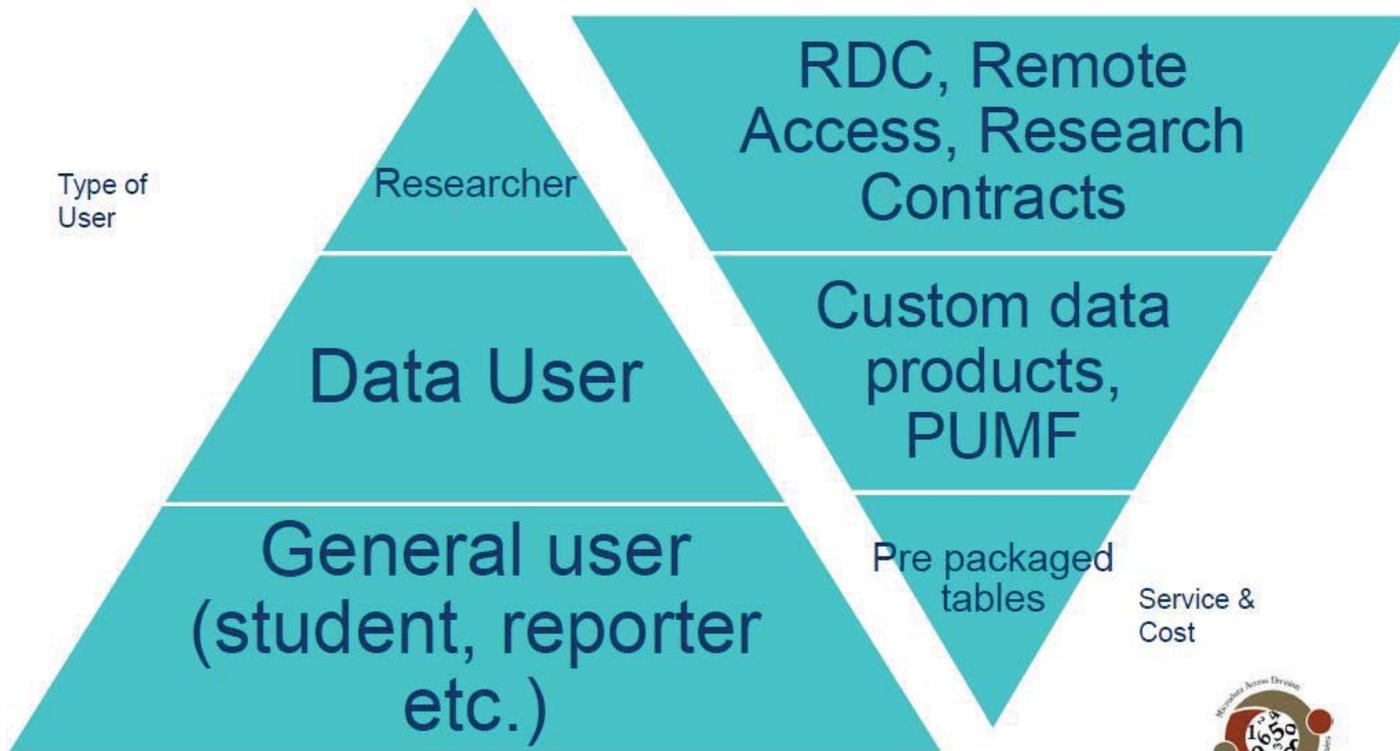
# Continuum of Access



Statistics Canada • Statistique Canada



# Users and costs



## Vision for Microdata Access at Statscan

- Microdata Access Division: One point of contact for external requests to access microdata
- Spectrum of access services
- Responsive, innovative division with expertise on protecting confidentiality
- Cost efficiencies through harmonization, resource sharing and minimizing duplication



## Principles for Microdata Access

- Access cannot be considered separately from the rest of the data process, or from other forms of access.
- Confidentiality must be protected
- Access provided will be for research, for the public good
- Access costs will be covered by the researcher or research community
- Multiple access options will be made available as part of a continuum

Statistics Canada • Statistique Canada



# What do Researchers Access?

## ■ Confidential Microdata Files

- For each survey conducted, a master file is constructed which contains all responses by each respondent, recorded in the format specified on the questionnaire.
- All obvious identifiers have been removed from the master files (i.e. Names, address, phone numbers etc.)
- Only two types of users are permitted to access the master file – the author division (to create extractions for paying clients and for divisional analysts to perform their research) and Research Data Centre (RDC) analysts. Please note: \*Not all master files are available in the RDCs.
- When analysis is conducted on master files, the results of the analysis must be vetted through a process called “disclosure analysis” to ensure that it conforms to the confidential rules established by Statistics Canada”. This is to ensure that no particular respondent is identifiable.



## What do Researchers Access?

- **Public use microdata files (PUMFs)**

- Each Public Use Microdata File is based on a corresponding master data file. The modifications performed by Statistics Canada before the PUMF is released ensure that the risk of breaching confidentiality has been removed. Since the results of any analysis performed do not have to be scrutinized before they are released, the file is considered “Public”.
- Modifications made to master files to convert them to PUMFs may include: collapsing of variables (e.g., age groups instead of individual years of age); collapsing variables into one variable (e.g., multiple language questions collapsed into one language variable for analysis); suppressing variables (although the variable is part of the master file, it will not show up in the public file); and removing outliers (removing cases that are extremes - often used with income).
- By using these techniques to anonymise the files, combining variables will not result in the user identifying a respondent.



## Managing Access

- Microdata Access Division
- Programs related to accessing PUMFs
  - Data Liberation Initiative
  - Access to PUMF collection
- Programs related to accessing microdata
  - Real Time Remote Access
  - Research Data Centres
  - Federal Research Data Centre
  - Canadian Centre for Data Development and Economic Research (CDER)



## Data Liberation Initiative

- Data Liberation Initiative (DLI): Post-secondary institute members access the entire collection of PUMFs, plus:
  - DLI metadata are coded for searchability;
  - DLI members have support through a very active listserv;
  - Annual training of data librarians takes place across Canada every year.
  - 75 subscribers currently



## Benefits of the DLI

- Data and network of researchers
  - 75 Data librarians across Canada accessing 350 PUMFs, aggregate data and geographic information files
- Data access infrastructures
  - Nesstar, Website, FTP
- Data expertise
  - Focal point experts
  - SM division experts
  - Networked community



# DLI Governance

- External:
  - Data Liberation Initiative External Advisory Committee
- Internal
  - Steering Committees
  - Microdata Access Management Committee
  - Microdata Release Committee
  - Policy on Microdata Access
  - Directive on Microdata Release



## Access to PUMF Collection

- Access to Statistics Canada's PUMF Collection:
  - Like the DLI, but for organizations who do not qualify as a post-secondary institute.
  - Similar structure and support as the DLI
  - International and Canadian membership



## Individual access to PUMF's

- Access to Statistics Canada's PUMF Collection:
  - Not limited to consortia
  - Individuals can sign a license with STC to use a PUMF for private or public uses.



# Accessing Confidential Microdata

## ■ Key differences to note

- PUMF's front load confidentiality (sacrificing precision for broader dissemination)
- No confidential information leaves Statscan
- Accessing confidential information requires much more accountability
- Costs increase dramatically

## Real Time Remote Access

- On-line remote access facility allowing users to run more or less in real-time, data analyses on microdata located in a central, secure location.

Tabular frequencies only - SAS only

Limited to particular household surveys data sets

Confidentiality automated

No primary manual intervention

Deemed employee status not required

Ability to interrogate RTRA micro data at any time\*

Access from any computer with internet access, using a secure username and password - No travel to RDCs

# Research Data Centre Program

## Research Data Centres

- Secure data laboratories located in 26 Universities across Canada and 1 research institute in Yellowknife.
- Centres are financed by the universities and funding councils, and are operated by Statistics Canada.
- Academic and government researchers apply to gain access to the data housed in the RDCs.
- Currently social and health survey data and some administrative health and justice data are housed in the RDCs.



## The Research Data Centre Program

It is a partnership that includes:

- More than 40 Canadian universities
- Major Granting Councils (SSHRC, CIHR, CFI)
- Provincial governments

The program comprises:

- Over 50 employees
- Over 2650 projects since inception with 770 active research projects
- Nearly 4000 researchers



## Access to the Research Data Centres (Academic, NGO and private sector researchers)

- Researcher submits a project proposal
- StatCan and peers evaluate proposal
- Researcher requires a security clearance - enhanced reliability check
- RDC analysts gives an orientation session at which time the “oath of office” and contract are signed
- Researcher agrees to provide publicly available report that falls within Statistics Canada’s mandate

# Evaluation Criteria

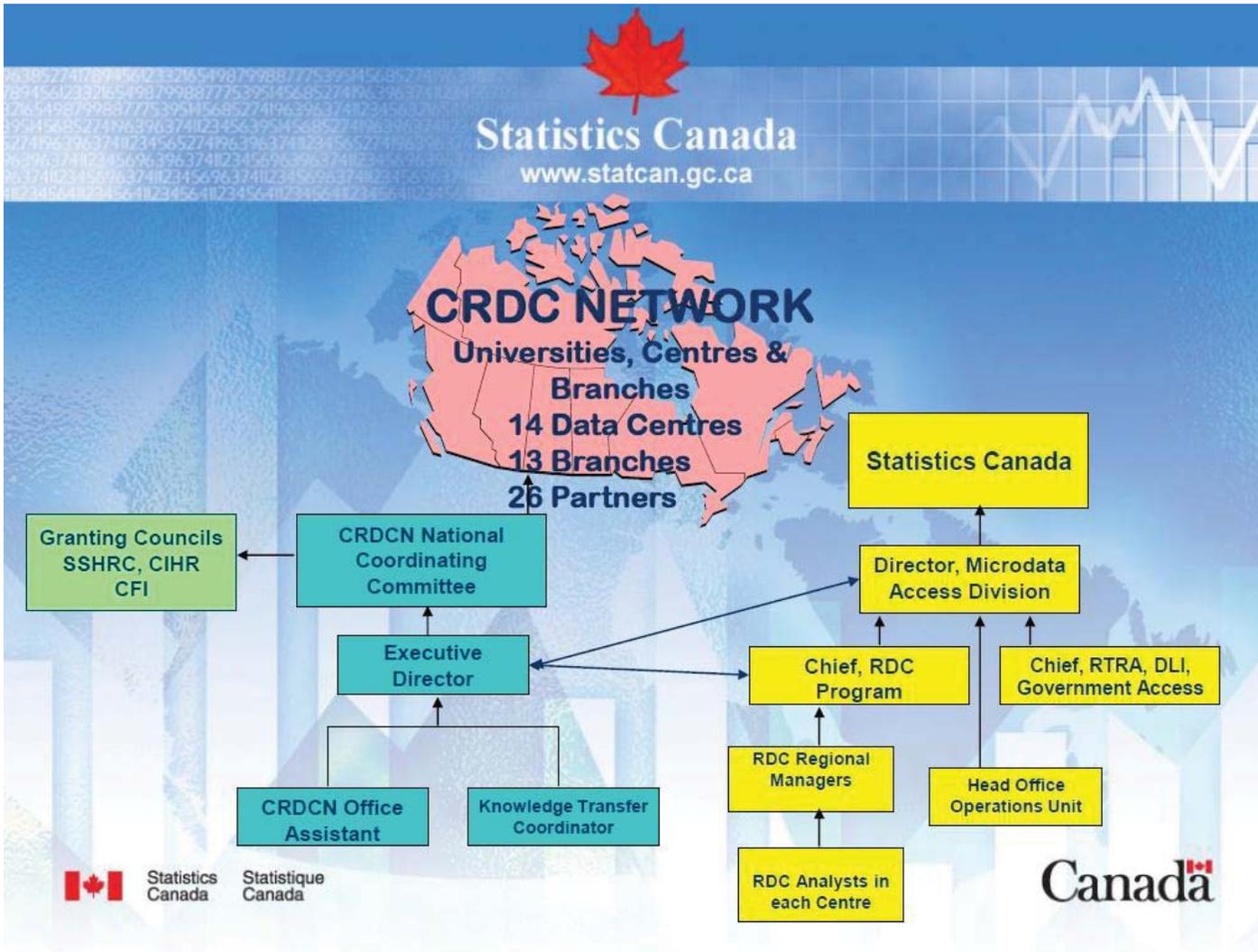
(academic, NGO or private sector researchers)

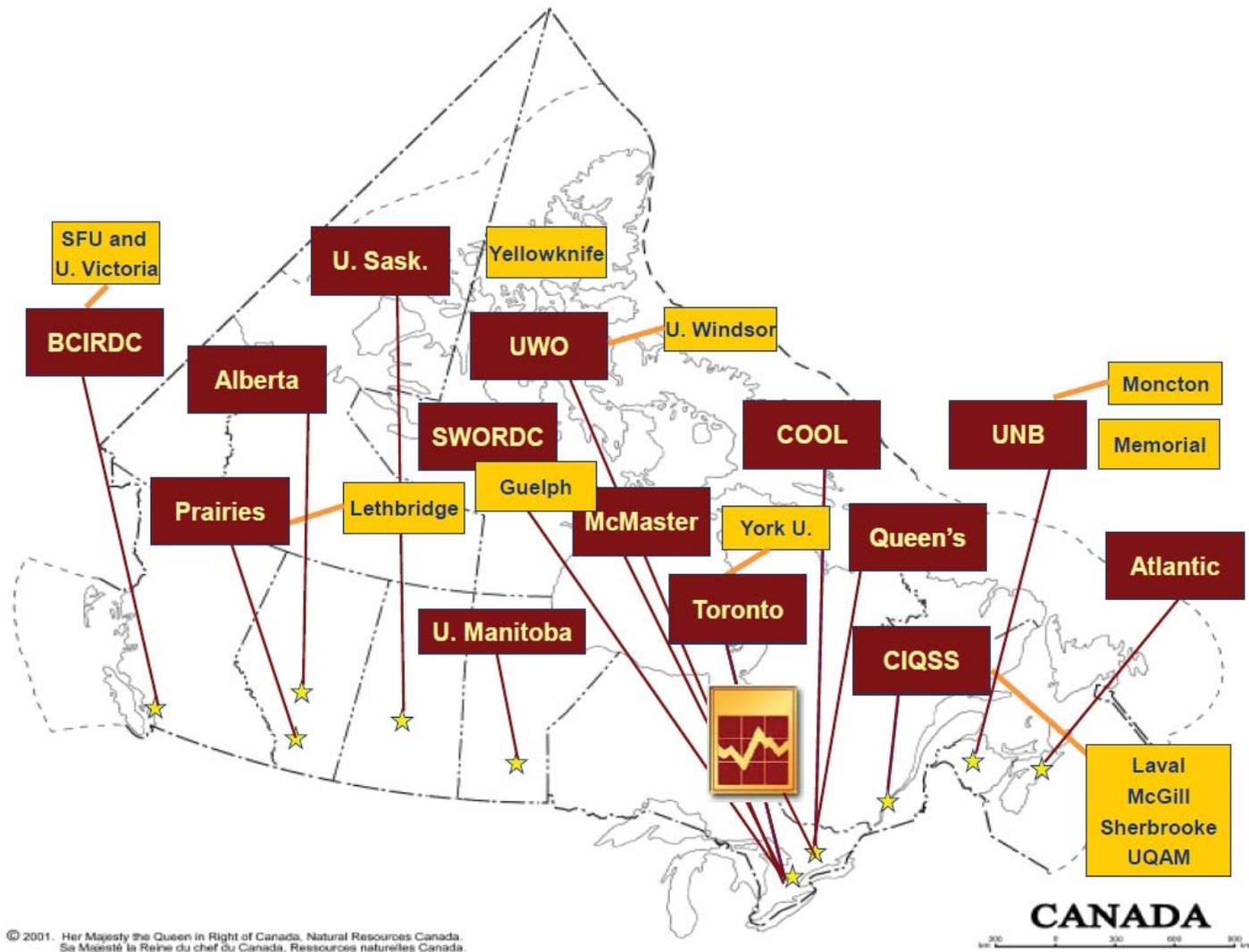
## StatCan review

- Determines a demonstrated need for access to detailed microdata
- Ensures that requested data can address the research questions
- Examines potential conflict of interests

## Peer Review

- Clearly defined project objectives
  - Suitability of analytical and statistical methods
  - Scientific merit of the project
  - Do the applicant and the team members (if applicable) have the experience, qualifications, and expertise to successfully complete the proposed project?
- Students require letter of support from supervisor





## Federal Research Data Centre

Federal Research Data Centre opened in 2003

Motivation for starting the program:

- Provide a secure site where federal government employees can conduct complex statistical analysis
- Meet the research demands and simplify the data access procedures for federal government employees
- Expand the research output from StatCan data holdings, while maintaining the confidentiality of information provided by Canadians



# Federal Research Data Centre

## Federal Research Data Centre:

- Located at Statistics Canada for Federal Government researchers
- Agreements approved by departmental focal points
- Purchase the use of one or more seats for a given period (up to 1 year at a time)
- Statistics Canada analysts provide support



## Canadian Centre for Data Development and Economic Research (CDER)

- The Canadian Centre for Data Development and Economic Research (CDER) is the repository for business and economic micro data files containing sufficient detail for researchers to undertake meaningful and complex analysis while at the same time safeguarding the confidentiality of the results and the privacy of business respondents and workers in these establishments.
- CDER is designed to strengthen how Statistics Canada organizes its business micro-economic data resources to support the information needs of Statistics Canada, other federal government departments and the research community. The Centre increases the capability of Statistics Canada to serve the policy community, brings together a critical mass of resources to do so, established processes to guarantee the security and confidentiality of the data, and will improve Statistics Canada's ability to meet its mandate in the area of analysis.



## Canadian Centre for Data Development and Economic Research

- Managed by Microeconomic Analysis Division
- Two stage review process feasibility and detailed project review/costing
- CDER works with data divisions to produce analytic data file
- Federal Departments and Academics.



## Canadian Centre for Data Development and Economic Research (CDER)

- Proposals must have enough detail so that the feasibility (from a data perspective).
- Individuals who wish to obtain custom tabulations should contact the appropriate subject matter division at Statistics Canada.
- The justification for the research including:
  - context;
  - the research question; and
  - the project's contribution to the literature.
- The analytical framework and data requirements including:
  - detailed data requirements;
  - proposed methodology;
  - justification for using the micro data;
  - the expected output (the number and type of tables);
  - software requirements; and
  - expected length of the project and start date.



## Data Files Accessible through CDER

- [Annual Survey of Manufacturing](#)
- [Annual Survey of Manufacturing – Export and Import Registry Database](#)
- [Canada Border Service Agency Customs Database](#)
- [Capital and Investment Program](#)
- [Longitudinal Employment Analysis Program](#)
- [Longitudinal Worker File](#)
- [National Accounts Longitudinal Microdata File](#)
- [T2-LEAP](#)
- [T2-LEAP-Export and Import Registry Database](#)
- [Survey of Financing of Small and Medium Enterprises](#)
- [Survey of Innovation and Business Strategies](#)
- [Workplace Employee Survey](#)



## How to contact us

- ICN address:
  - [http://icn-rci.statcan.ca/rdc/rdc\\_r000\\_e.htm](http://icn-rci.statcan.ca/rdc/rdc_r000_e.htm)
- Statistics Canada Website address:
  - <http://www.statcan.gc.ca/rdc-cdr/index-eng.htm>
  - <http://www.statcan.gc.ca/dli-ild/dli-idd-eng.htm>
- Canadian Research Data Centre Network Website address:
  - <http://www.rdc-cdr.ca>



## 附属資料Ⅲ 参考ウェブサイト・文献一覧

### 国内事例関連

#### 統計局

<a href="http://www.stat.go.jp/data/kokusei/2010/users-g/pdf/tyousa.pdf">http://www.stat.go.jp/data/kokusei/2010/users-g/pdf/tyousa.pdf</a>	平成 22 年国勢調査調査票
<a href="http://www.stat.go.jp/info/guide/public/keikaku/pdf/opendata.pdf">http://www.stat.go.jp/info/guide/public/keikaku/pdf/opendata.pdf</a>	統計におけるオープンデータの高度化について

#### 統計センター

<a href="https://www.e-stat.go.jp/SG4/onlinet/guide/help.html">https://www.e-stat.go.jp/SG4/onlinet/guide/help.html</a>	政府統計オンライン調査システムの画面解説
<a href="http://www.nstac.go.jp/services/society_paper/25_05_04_PowerPoint.pdf">http://www.nstac.go.jp/services/society_paper/25_05_04_PowerPoint.pdf</a>	統計データの二次的利用に関する統計センターの取組状況

### 海外事例関連

#### アメリカセンサス局(United States Census Bureau)

<a href="https://www.census.gov/">https://www.census.gov/</a>	センサス局トップページ
<i>The Use of Electronic Reporting and Other Electronic Tools for Data Providers</i> <a href="http://www.amstat.org/meetings/ices/2012/papers/302090.pdf">http://www.amstat.org/meetings/ices/2012/papers/302090.pdf</a>	センサス局におけるインターネット回答に関する取組(Centurion)について
<a href="https://www.census.gov/dataviz/">https://www.census.gov/dataviz/</a>	データ視覚化について

#### アメリカ労働統計局(United States Bureau of Labor Statistics)

<a href="http://www.bls.gov/ces/">http://www.bls.gov/ces/</a>	CES に関するトップページ
<a href="http://www.bls.gov/opub/hom/">http://www.bls.gov/opub/hom/</a>	BLS における調査手法について
Data Collection in the U.S. Bureau of Labor Statistics' Current Employment Statistics Survey <a href="http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.44/2012/mtg2/WP20.pdf">http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.44/2012/mtg2/WP20.pdf</a>	CES における調査手法について

#### カナダ統計局(Statistics Canada)

<a href="http://www23.statcan.gc.ca/imdb-bmdi/pub/3701-eng.htm">http://www23.statcan.gc.ca/imdb-bmdi/pub/3701-eng.htm</a>	LFS に関するトップページ
<i>Guide to the Labour Force Survey</i> <a href="http://www.statcan.gc.ca/pub/71-543-g/71-543-g2012001-eng.pdf">http://www.statcan.gc.ca/pub/71-543-g/71-543-g2012001-eng.pdf</a>	調査内容・回答方法などに関する解説

#### その他

統計データの二次利用促進に関する研究会 報告書(平成 20 年 8 月)	諸外国におけるデータ・アーカイブ、マイクロデータの利活用状況に関する報告書
統計データの二次的利用促進に関する研究会 平成 23 年度報告書(平成 24 年 7 月)	
<a href="http://www.icpsr.umich.edu/">http://www.icpsr.umich.edu/</a>	ICPSR に関するトップページ
<a href="http://www.dataenclave.org/">http://www.dataenclave.org/</a>	Dataenclave(NORC)に関するトップページ