

## 제3장

# 마이크로데이터 제공 현황에 관한 해외사례 연구

김경미 · 이의규 · 정미옥

## 제1절 서론

### 1. 마이크로데이터 제공의 개요

통계자료가 적시적지에 활용될 수 있도록 그 수요에 올바르게 대응하여 제공하는 것은 정확한 통계를 생산해 내는 것만큼이나 중요한 일이다. 통계자료는 일반적으로 원시자료(Raw data), 마이크로데이터(Microdata), 그리고 매크로데이터(Macrodata)로 구분된다. 통계청 MDSS(Microdata Service System)의 용어 정의에 따르면, 원시자료(Raw data)는 입력 첫 단계의 자료로서 입력오류나 조사오류 등이 걸러지기 이전 단계의 자료를 말하며, 마이크로데이터(Microdata)는 오류를 수정한 자료로서 통계표 작성 등 데이터 가공의 바탕이 되는 자료를 의미한다. 또한 매크로데이터(Macrodata)는 마이크로데이터를 임의의 기준에 따라 집계한 자료를 지칭하는데, 집계의 정도에 따라 세분된 자료부터 통합된 자료까지 다양하게 제공된다.

이용자에게 제공되는 데이터는 주로 마이크로데이터와 매크로데이터이다. 일반적으로 매크로데이터가 제공되는 형식은 통계표(Tables) 형태로 가공하여 서비스되는 형식으로 통계청에서는 간행물이나 KOSIS

(Korean Statistical Information Service)를 통해 제공되고 있다. 반면 마이크로데이터가 제공되는 형식은 외부이용자를 위해 사전에 만들어진 마이크로데이터 파일로서 CD-ROM 등에 저장하여 제공하는 공공이용파일(Public Use File)의 형태, 특정한 심층이용자에게 제공되며 제공 전에 법적인 조치나 서명 등이 필요한 인가파일(Licensed File)의 형태, 그리고 통신망을 통해 제공하는 원격접속파일(Remote Access Files)의 형태가 있다. 이 밖에 통계기관의 엄격한 감독하에 기관 내에서 마이크로데이터를 이용하는 제도로 데이터실험실(Data Laboratories)이 있다. 이상의 데이터 제공형식을 정리하면 <표 3-1>과 같다.

<표 3-1> 데이터 제공 형식

형 식	자료의 특징 및 내용	통계청 적용형태	
통계표(Tables) 형태로 가공하여 서비스	· 심층 이용자 요구에 따라 통계표 작성 · 지나치게 세분화되면 비밀노출 위험	간행물 KOSIS	
익명자료 파일 (AMF <sup>1)</sup> )	공공이용 파일(Public Use File)	· 외부이용자를 위하여 사전에 만들어진 마이크로데이터 파일로 CD-ROM 등에 저장	마이크로 데이터 CD-ROM
	인가파일 (Licensed File)	· 특정한 심층이용자에게 제공되며, 제공 전에 법적인 조치(서명 등)가 필요	개별자료 제공
원격접속파일 (RAF <sup>2)</sup> )	· 통신망을 통하여 심층 이용자들이 마이크로데이터를 활용하는 제도	MDSS	
데이터실험실 (DL <sup>3)</sup> )	· 통계기관의 엄격한 감독하에 기관 내에서 마이크로데이터를 이용하는 제도	On-site	

주: 1) Anonymised Microdata Files.

2) Remote Access Files.

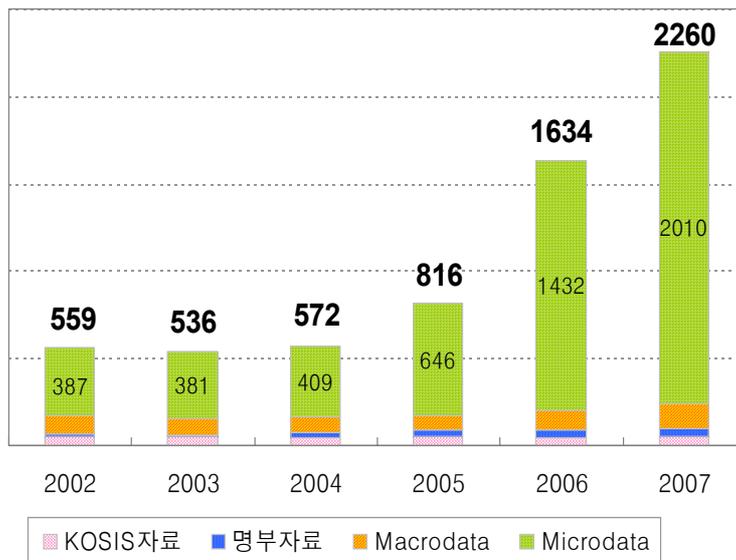
3) Data Laboratories.

자료: 박원환·황조연, 「통계자료의 특성과 비밀보호 방법에 관한 연구」, 2004; 통계청, “통계마이크로데이터 제공현황 및 개선방안”, 2006.

## 2. 연구의 목적 및 필요성

사회구조가 복잡해짐에 따라 의사결정의 기초자료로서 통계수요가 급격히 증가하고 있다. 특히, 중앙통계기관인 통계청의 자료는 각종 국가정책 수립 및 평가, 기업경영, 사회현상 탐구 등의 기초자료로 활용되는데, [그림 3-1]과 같이 2002년 559건의 자료가 이용되었음에 비해 2007년에는 2,260개의 자료가 이용되어 그 수요가 급격히 증가함을 알 수 있다.

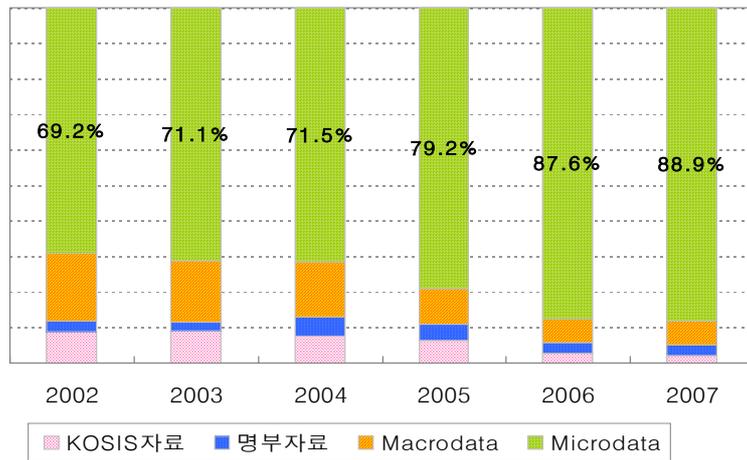
[그림 3-1] 통계청 제공자료수



주: 2002~04년은 건수, 2005~07년은 개수.  
자료: 통계청(2007).

특히 초기 단순한 집계자료 요청에서 점차 다양한 분석이 가능한 마이크로데이터의 요청이 증가하고 있는데, [그림 3-2]를 보면 제공되는 전체 자료 중 마이크로데이터 비중이 점점 증가하고 있음을 알 수 있다.

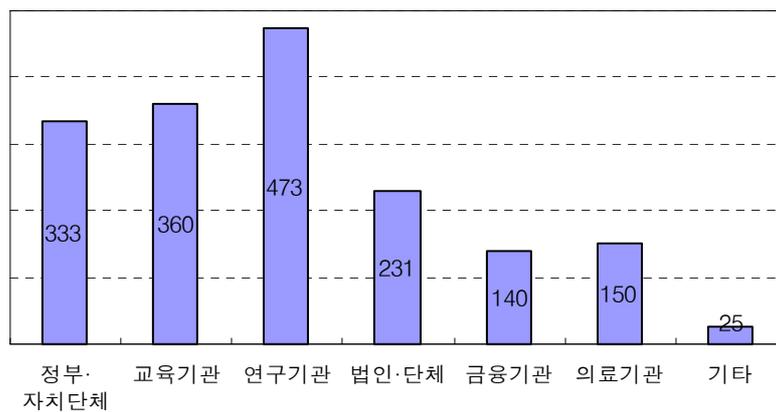
[그림 3-2] 통계청 제공자료 구성비



주: 2002~04년은 건수, 2005~07년은 개수.  
 자료: 통계청(2007).

또한 제공되는 통계자료는 다양한 기관에서 이용되는데 2007년 현재 [그림 3-3]과 같이 연구기관, 교육기관, 정부·자치단체, 법인·단체의 순으로 많이 이용되고 있는 것으로 나타났다.

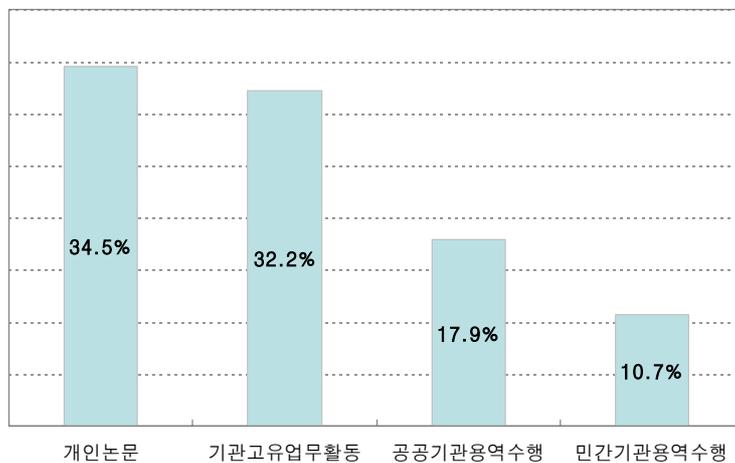
[그림 3-3] 요청기관별 자료제공 현황(2007)



자료: 통계청(2007).

2006년 통계청 통계정보국에서 실시한 “마이크로데이터 활용도 제고를 위한 심층의견조사” 결과에 따르면 [그림 3-4]와 같이 그 이용목적 또한 다양하다. [그림 3-4]에서 보듯이 개인논문과 기관 고유 업무활동 부분이 약 2/3를 차지했으며, 공공기관 및 민간기관의 용역수행을 위해 이용하는 경우도 일정 비율을 차지하였다.

[그림 3-4] 마이크로데이터의 이용목적



자료: 통계청, “마이크로데이터 활용도 제고를 위한 심층 의견조사”, 2006.

이와 같이 마이크로데이터의 이용이 증가함에 따라 마이크로데이터의 제공방법 및 비용 등에 관한 의견들도 분분하다. 또한 이용기관과 이용목적도 다양하므로 그에 적합한 제공방법에 대해 연구할 필요가 있는바, 해외의 마이크로데이터 제공 현황에 관한 사례를 자세히 파악함으로써 제도적 또는 방법론적 측면에서 적용방안을 모색하고자 한다.

다음 제2절에서는 우리나라 통계청의 마이크로데이터 제공현황에 대해 살펴보고, 이후 제3절에서 미국, 캐나다, 호주, 영국, 그리고 일본의 현황을 각각 파악하여 마지막 절에서는 이들로부터 시사점을 도출한다.

## 제2절 통계청 현황

### 1. 제공방법

해외의 현황을 살펴보기에 앞서, 우리나라 통계청에서는 어떻게 마이크로데이터를 제공하고 있는지를 알아보려고 한다. 마이크로데이터는 크게 세 가지 방법으로 이용자에게 제공된다. 먼저 개인 식별 데이터를 제거한 마이크로데이터를 CD-ROM에 저장하여 제공하는 형식이 있다. 이는 통계쇼핑몰(<http://www.nso.go.kr/shop>)을 통해 제공되고 있다. 둘째로 통신망을 이용한 방법으로 이용자가 인터넷을 통해 시스템에 접속하여 원하는 자료를 직접 가공하거나 다운받을 수 있도록 지원하는 윈-스톱 서비스 시스템인 MDSS(MicroData Service System)를 이용하는 방법이다. MDSS는 2006년 1월 1일부터 서비스를 시작하였고 홈페이지를 통해 제공 가능한 자료의 목록과 범위, 이용방법, 비용 등을 안내하고 있으며, MDSS 사이트(<http://mdss.nso.go.kr>)에 접속하여 신청한 후 자료를 다운로드 받을 수 있다. 중앙정부 및 지방자치단체에서 신청할 경우에는 통계청에서 작성·제공하고, 대학 및 연구기관에서 신청할 경우에는 통계협회로 이첩하여 제공하는 방식을 취하고 있다. 마지막으로 특정장소에서 좀 더 심층적인 데이터를 제공받을 수 있는 On-site Access 방법이 있다. 현재 2개소를 운영하고 있는데 통계청(대전)과 대한통계협회(서울) 내에서 운영하고 있다.

### 2. 제공자료

통계청에서 작성되는 통계는 조사통계가 44종, 가공통계가 11종, 보고통계가 1종으로 총 56종의 통계가 작성되고 있다. 이 중 마이크로데이터가 제공되는 통계는 조사통계 32종과 기업체모집단명부를 포함하여 총 33종에 대해 제공하고 있다. 제공되는 조사통계 목록은 <표 3-2>와 같다.

〈표 3-2〉 통계청 제공통계 목록

No.	제공되는 조사통계	No.	제공되는 조사통계
1	가계조사	16	사회통계조사
2	가계자산조사	17	산업총조사
3	가구소비실태조사	18	생활시간조사
4	건설업통계조사	19	서비스업총조사
5	경제활동인구조사	20	양곡소비량조사
6	광업제조업통계조사	21	어가경제조사
7	국부통계조사	22	어업기본통계조사
8	기업체모집단명부	23	어업총조사
9	농가경제조사	24	운수업통계조사
10	농산물생산비조사	25	인구동태조사
11	농어업법인사업체통계조사	26	인구이동통계
12	농업기본통계조사	27	인구주택총조사
13	농업총조사	28	인력실태조사
14	도소매업 및 서비스업통계조사	29	임업총조사
15	사망원인통계	30	전국사업체기초통계조사

주: 1) 인구총조사와 주택총조사, 도소매업 통계와 서비스업 통계는 통합해 총 30종의 통계로 제공.

2) 2005년 인구주택총조사의 경우는 1%, 2%, 5% 표본자료를 제공하고 있음.

각 조사별 자료의 제공범위는 자료제공심의회를 거쳐서 결정되며 결정된 제공범위는 MDSS 홈페이지에 게시되어 있다. 자료제공심의회는 통계청 통계정보국장이 위원장이 되며 통계작성 담당과장 등 10인의 위원이 모두 통계청 내부자로 구성된다. 자료제공심의회에서 주로 심의하는 내용은 ① 통계조사별 자료제공범위 설정에 관한 사항, ② 통계목적으로 사용여부의 판단에 관한 사항, ③ 제한적 공개가능 자료에 대한 제공여부에 관한 사항, ④ 자료제공업무의 효율성 향상과 ⑤ 이용자의 편의제고를 위한 제도개선 및 운영방법에 관한 사항이다. 자료의 제공범위를 심의함에 있어서는 응답단위의 비밀보호와 응답단위·항목별 자료의 신뢰성 등에 중점을 둔다. 심의회는 자료제공범위의 변경 필요성 등이 있을 경우 수시로 개최하여 심의한다.

### 3. 가격책정

마이크로데이터를 중앙정부와 지방자치단체에서 요청할 경우 무상으로 제공하며, 대학, 연구기관 및 개인이 요청할 경우 유상으로 제공한다. 마이크로데이터의 비용 산정 기준은 수익자부담원칙<sup>1)</sup> 및 실비보상 원칙을 바탕으로 통계청통계자료규정 제5조 제3항의 규정에 의하여 정하고 있다. 자료이용료는 1MB에 기본료 20,000원이 부과되며 초과되는 1MB당 1,500원이 추가된다. 예를 들어 약 12만 개의 전국의 광업·제조업 사업체에 대해 조사되는 광업·제조업통계조사 전체자료가 130MB인 경우, 기본료 20,000원과 이를 초과하는 129MB에 대해 1MB당 각 1,500원이 추가되어 총 213,500원<sup>2)</sup>이 부과된다.

주문형 위탁처리의 경우는 자료이용료와 별도로 소프트웨어개발비가 부과되는데 이는 정보통신부 고시 제2005-22호 「소프트웨어사업 대가의 기준<sup>3)</sup>」을 준용하여 징수한다. 소프트웨어개발비가 추가되는 인구주택총조사의 표본추출의 경우 기본추출틀은 조사구 100개, 조사구당 특성치 2개이며 그 산출내역은 자료이용료(기본 1MB = 20,000원)에 소프트웨어개발비 198,010원이 더해져 218,010원이 된다. 기본특성치(2개)를 초과할 경우 특성치 1개당 37,120원이 추가 부과되며 특성치가 기본(2개)인 상태에서 기본조사구(100개)를 초과할 경우 초과조사구 100개당 기본사용료 218,010원이 추가 부과된다.

MDSS로 제공되는 경우는 마이크로데이터 비용산정 기준이 그대로 적용되며, CD-ROM으로 제공되는 경우는 마이크로데이터 비용산정 기준에 수수료 2%가 부과된다. On-site Access의 경우 같은 기준에 이용환

---

1) 일반적으로 공공사업의 실시에 의하여 특별히 이익을 받는 사람에게 소요경비의 전부 또는 일부를 부담시키는 일을 의미하며, 마이크로데이터를 서비스하기 위해서는 별도의 마이크로데이터 관리, 전산처리 등의 통계청의 자원이 추가로 소요되기 때문에 추가로 소요되는 자원에 대한 비용을 이용자에게 부담시키고 있음.

2) 20,000원 + 1,500원 × 129MB = 213,500원.

3) 소프트웨어개발비 = 개발원가 + 직접경비 + 이윤.

· 개발원가 = 보정전개발원가(코드라인수 × 코드라인당 단가) × 보정계수.

· 직접경비 = 컴퓨터시스템 사용료, 소프트웨어 도구 사용료, 제출문서의 인쇄, 청사진비 등 특별히 소요되는 직접비용.

· 이윤 = 개발원가 × 10%.

경 설정, 시스템 이용 등에 대한 수수료 10%가 추가로 부과된다.<sup>4)</sup>

#### 4. 자료신청절차

통계청의 마이크로데이터 이용은 공표통계에서 제공할 수 없는 보다 심층적인 사회현상 분석을 원하는 다양한 계층의 이용자에게 마이크로데이터를 제한 없이 제공함을 원칙으로 한다. 외국기관을 포함한 외국인에 대한 제공은 국제기구, 외국정부기관 등 신뢰성 있는 기관에 한정하여 국내 기준과 동일하게 적용함을 원칙으로 하고, 사안별로 담당과에서 검토를 하고 필요시 자료제공심의회에서 심의한다.

MDSS를 이용해 사용자가 직접 추출하여 다운로드 받고자 하는 경우는 홈페이지에서 회원가입을 하여야 한다. KOSIS에서 무료로 제공되는 자료의 목록 확인과 이용할 자료의 자료제공 범위를 확인 후 자료조회 및 필요한 항목을 선택하고 승인을 요청한다. 자료제공 유무가 결정되면 비용을 납부하고 자료를 다운로드 받을 수 있다. 통계자료를 위탁처리하는 경우는 자료제공 범위 확인까지는 동일한 절차이며, 그 후 위탁신청서를 작성하여 신청하면 통계청에서 자료제공유무를 결정하여 자료를 가공하고 이용자가 비용을 납부하면 자료가 제공된다. On-site Access의 경우에는 이용을 승인받기 위해서 On-site Access 이용신청서 및 연구계획서를 같이 제출해야 한다. 이 연구계획서에서는 통계청에서 제공하는 통계조사보고서, KOSIS 등의 공표자료, 그리고 기존의 마이크로데이터 상품으로는 자료이용 목적을 충족할 수 없음을 입증하여야 한다. 이용을 승인받은 후에는 통계청 또는 대한통계협회 내의 지정장소에서 사전 설치되어 있는 전산장비를 이용하여 마이크로데이터 분석 작업을 한다. 또한 자료이용자의 분석결과에 대하여, 자료생산부서에서 응답단위의 비밀보호의무 준수 여부 및 분석 결과의 통계적 신뢰도 등을 검토 후 외부 반출 여부를 결정한다. 외부로 분석결과를 반출을 결정한 경우 자료이용자는 비용을 납부하고 자료를 제공받을 수 있다.

4) 현재 총괄원가보상원칙으로 비용산정 기준의 개정을 검토 중에 있음.

## 5. 제공 후 관리

통계청은 자료 이용시 지켜야 할 사항을 통계자료 제공규정 제7조 제2항으로 정하고 있다. 포함되는 사항은 개별자료에 의거 알게 된 사항에 대한 누설금지, 통계목적 이외의 사용금지, 제공 자료의 복제 및 대여금지, 자료 사용 후 즉시 폐기이다. 이러한 의무를 위반할 시에는 개정통계법<sup>5)</sup> 제41조 3항에 의거하여 100만원 이하의 벌금이 부과된다.

CD-ROM이나 MDSS로 제공되는 경우 마이크로데이터를 이용한 분석결과에 대한 검토는 이루어지지 않고 있으며, On-site Access의 경우에는 분석결과의 반출 전 응답단위의 비밀보호의무 준수 여부 및 분석 결과의 통계적 신뢰도 등을 검토한 후 반출여부를 결정한다.

## 제3절 국외 현황

국외의 마이크로데이터 제공현황은 각국의 통계청과 마이크로데이터를 제공하는 각 기관의 홈페이지, 관련 문헌들을 참조하였고 각 국 통계청의 자료제공 담당자와의 이메일을 통해 자료를 수집하였다. 또한 미국과 캐나다의 On-site방법과 미국 노동통계국의 off-site방법은 많은 부분 神林 龍(2007)의 보고서를 발취 인용하여 정리하였음을 밝혀둔다.

### 1. 미국 센서스국(Census Bureau)

#### 가. 공공이용파일(Public Use File)

미국은 공공이용파일(Public Use File)<sup>6)</sup>을 다양한 기관에서 간단한 절차를 통해 거의 무료로 제공하고 있다. 센서스국은 American Factfinder를 통해 직접 제공하고 있으며 웹사이트(<http://factfinder.census.gov>)에서

5) 개정통계법은 2007년 10월 28일자로 시행된다.

6) 외부이용자를 위하여 사전에 만들어진 마이크로데이터 파일로, 해외의 경우 비밀보호 기법이 적용된 공공파일이다.

다운로드 할 수 있다. 약간의 실비를 지급하면 CD-ROM으로 구매도 가능하다. 또한 Data Ferrett 사이트(<http://dataferrett.census.gov>)에서 다양한 자료를 제공받을 수 있는데 이메일주소를 기입한 후 무료로 다운로드 받을 수 있다. 이 사이트에서는 마이크로데이터의 제공뿐만 아니라 사이트 내에서 프로그램을 이용하여 집계표 작성 및 그래프 작성이 가능하다. 제공되는 자료는 총 20종이며 그 목록은 <표 3-1>과 같다.

<표 3-3> Data Ferrett 제공자료 목록

No.	제공되는 통계자료
1	American Community Survey (ACS)
2	American Housing Survey (AHS)
3	Behavioral Risk Factor Surveillance System (BRFSS)
4	Consumer Expenditure Survey (CES)
5	County Business Patterns (CBP)
6	Current Population Survey (CPS)
7	Decennial Census of Population and Housing (Census 2000)
8	Decennial Census of Population and Housing (Census 1990)
9	Delaware Statistics
10	Harvard MIT data Center Collection
11	Home Mortgage Disclosure Act (HMDA)
12	Maryland Statistics
13	National Ambulatory Medical Care Survey (NAMCS)
14	National Health and Nutrition Examination Survey (NHANES)
15	National Health Interview Survey (NHIS)
16	National Survey of Fishing, Hunting, and Wildlife-Associated Recreation (FHWAR)
17	Small Area Income and Program Poverty Estimates (SAIPE)
18	Social Security Administration
19	Survey of Income and Program Participation (SIPP)
20	Survey of Program Dynamics (SPD)

자료: Data Ferrett 홈페이지.

미네소타 인구센터에서 운영하는 IPUMS(Integrated Public Use Microdata Series)에서는 주로 센서스 자료를 제공하고 있는데 간략한 연구계획서 제출을 통해 회원으로 가입한 후 무료로 다운로드 받을 수 있다. 제공 절차를 자세히 살펴보면, 먼저 사이트(<http://www.ipums.org>)에서 사용자 등록 후 로그인을 하고 자료 추출 요청을 한다. 이때 국가, 연도, 대상집단, 변수를 지정하면 미네소타 인구센터에서 자료를 추출한 후(수분에서 수시간 소요), 연구자에게 이메일로 통보하게 되는데 72시간 내에 자료를 다운로드 받아 이용할 수 있다. 자료는 128비트 암호화 장치에 의해서 전송된다. IPUMS는 미국의 센서스 자료뿐만 아니라 다른 나라의 센서스 자료도 제공하고 있다. 제공하고 있는 자료를 살펴보면, 1850년에서 2000년까지의 미국 센서스 자료, 미국의 CPS(Current Population Survey)자료, 국가별 센서스 자료(26개국, 80개의 센서스자료, 2억 2백만 개의 레코드) 등이 있다.

NBER(National Bureau of Economic Research)에서도 홈페이지(<http://www.nber.org/data>)를 통해 다운로드 받을 수 있도록 하고 있는데 산업별 생산성자료 등 각종 경제연구에 필요한 자료를 제공한다. 가공된 자료로는 개별 학자들이 취합하여 통합한 자료 중 검증이 된 자료를 제공하고 있다. 이들 자료는 무료로 제공되고 있으며 CPS자료를 저장매체로 요청할 경우 약 100달러의 실비가 부과된다.

CEPR(Center for Economic and Policy Research)에서는 CPS자료와 SIPP(U.S. Census Survey of Income and Program Participation)자료, 그리고 1980년에서 2000년까지의 센서스 1% 자료를 홈페이지(<http://www.ceprdata.org>)에서 제공하고 있다. 역시 무료로 제공되며 CPS자료를 저장매체로 요청할 경우 약 20~30달러의 실비가 부과된다.

Unicon Research라는 연구기관을 통해서도 제공되는데 제공되는 자료는 CPS자료, SIPP자료, NLTC(S(National Long Term Care Survey)자료, CPS자료에서 추출한 MDU(Medicare Data Utilities)자료가 있다. 홈페이지(<http://www.unicon.com>)의 회원가입 후 제공받을 수 있으며 100~1,500달러의 비용이 부과된다.

각 기관에서 제공되고 있는 센서스 자료는 다양한 형태로 제공되고 있는데, 2000년 센서스 자료의 경우는 1% PUMS(Public Use Microdata

Sample)자료, 5% PUMS자료, 괌 PUMS 10% 자료, 버진 아일랜드 PUMS 10% 자료, Long Form 전체자료(전체 조사대상의 1/6)가 제공되고 있다. 5% PUMS자료는 주와 10만 명 이상이 있는 카운티의 정보를 제공하며, 1% PUMS자료의 경우에는 표본 수는 작고 지리적(geographic) 변수는 40만 명 이상이 거주하는 지역의 아이디만 제공할 정도로 세분화 정도는 낮지만 지리적 변수를 제외한 다른 변수는 세분화된 자료를 제공한다. 5% PUMS자료가 어떤 인종이든 1만 명 이상의 인종을 대표할 경우 기타로 분류되어 있는 반면 1% PUMS자료는 자세한 인종 카테고리를 모두 공개한다. Long Form 전체자료의 경우는 데이터실험실(Data Laboratories: DL) 형식인 usRDC로만 이용가능하며 다른 방법으로는 접근이 불가능하다. 1990년 자료는 1% PUMS자료, 5% PUMS자료, 3% 노년인구자료, 0.5% 노동시장구역자료, 푸에르토리코 PUMS 5%가 제공된다.

#### 나. On-site(Data Laboratories)

##### 1) 개요

센서스국 내부의 경제 분석 부문인 CES(Center for Economics Studies)에서는 데이터 실험실(Data Laboratories) 형식의 usRDC(U.S. Research Data Center)를 운영하고 있는데, 이는 특정한 부지(site) 내에 설치된 특정한 단말기를 이용하여 데이터에 접근하고 거기에서 모든 작업을 완결하는 형식이다. 1994년 1월 보스턴에서 최초 개소한 이래 현재는 미국 전역의 9곳에서 운영하고 있다. 각각의 RDC(Research Data Center)는 센서스국에 있는 Census Bureau RDC, NBER(National Bureau of Economic Research)에 위치한 Boston Census RDC, UC Berkeley에 위치한 California Census RDC-Berkeley, UCLA에 위치한 California Census RDC-Los Angeles, FRBC(Federal Reserve Bank of Chicago)에 위치한 Chicago Census RDC, 미시건 대학에 위치한 Michigan Census RDC, Baruch School of Public Affairs에 위치한 New York Census RDC-Baruch, 코넬 대학에 위치한 New York Census RDC-Cornell, 듀크 대학에 위치한 Triangle Census RDC가 있으며 그 위치는 CES 홈페이지에 상세히 나타나 있다.

<표 3-4> 주요 지역 RDC 개요(2006년 10월 기준)

설치도시 (설치기관)	Boston (NBER <sup>1)</sup> )	LA (UCLA)	Chicago (FRBC <sup>2)</sup> )
설립년	1994년	1999년	2002년
관계 연구기관	NBER	UCLA	FRBC, Chicago Univ., Illinois Univ., Northwestern Univ., Argonne National Laboratory
센서스국 직원 수	1명 (경제학박사)	2명 (경제학박사 1명 포함)	
직원 외 스태프의 수	0명	0명	
부스의 수	8개	단말기 8개	
이용시간	24시간	24시간	
이용요금	\$15,000/년	\$3,125/월 \$15,000/년 소속대학원생 무료	
연간수락 proposal 수	12개		
이용자 중 대학원생 비율	약 50%	약 33%	

설치도시 (설치기관)	Ann Arbor (미시건대)	Ithaca NY (코넬대)	Durham (듀크대)
설립년	2002년	2006년	2000년
관계 연구기관	미시건대, ICPSR <sup>3)</sup>	Baruch Coll., NY city Univ., FRB <sup>4)</sup> NY, Fordham Univ., NBER, NY Univ., Pace Univ., Princeton Univ., Russell Sage Foundation, Rutgers Univ., NY State Univ., Yale Univ.	Duke Univ., M Carolina Univ., NC State Univ.
센서스국 직원 수	1명 (경제학박사)		1명 (경제학박사)
직원 외 스태프의 수	2명 (미시건대 교원)		1명 (듀크대 교수)
부스의 수	8개		8개
이용시간	24시간		24시간
이용요금	\$15,000/년 미시건대 교원, 대학원생 무료	비참가기관의 경우 \$20,000/년	\$15,000/년 관계대학 교원, 대학원생 무료
연간수락 proposal 수	15개		5~6개
이용자 중 대학원생 비율	약 50%		약 42%

주: 1) National Bureau of Economic Research. 2) Federal Reserve Bank of Chicago.  
3) Inter-University Consortium for Political & Social Research.  
4) Federal Reserve Bank.

자료: 神林 龍, 「北米における政府統計個票調査公開の現状に関する調査報告: 米國勞働統計局, 米國センサス局およびカナダ統計局のオンサイトリサーチを中心に」, 2007을 참조하여 정리.

usRDC에서 제공하는 자료의 목록은 <부록 A>에 첨부하였으며, 주요 지역 RDC의 개요는 <표 3-4>에 요약하였다. 마이크로데이터 이용시 부과되는 비용은 usRDC 관계기관의 이용자에게는 대부분 무료이나 관계 기관에 속하지 않는 이용자는 1년에 약 15,000달러를 지불해야 한다.

## 2) 자료신청 절차

RDC를 통한 자료의 신청까지는 다소 까다로운 절차가 요구된다. 먼저 웹사이트 상에서 이용하고 싶은 데이터를 확인 후 이용자의 계정(account)을 만들고 가장 가까운 usRDC에 체류하고 있는 센서스국 직원과 접촉한다. 이후 제안서(proposal)와 이력서 등의 필요서류를 첨부하여 온라인으로 제출하면 자료의 신청이 시작된다. 이용자의 요건은 국적이거나 소속의 제약은 없으나 usRDC를 이용하기 위해서는 SSS(Special Sworn Status)를 얻고, 무급이지만 연방직원으로서의 고용계약을 체결해야 한다.

제안서가 나타내야 할 기준은 ① 연구 내용이 연방법 13조에 규정된 센서스국의 직무내용을 채우는 것, ② 과학적으로 이점이 있는 것, ③ 공공이용파일(Public Use File)의 데이터에서는 분석 불가능한 것, ④ 분석의 성공 전망이 있는 것, ⑤ 비밀보호 정보·성과물의 취급에 관한 제한을 인정하는 것으로 정하고 있다. 제안서는 수시로 접수가 가능하며 제안서의 심사 프로세스는 다음과 같다. 우선 제출된 제안서에 대해 usRDC의 심사위원회 내에서 보통 2명의 담당자가 결정되어 6주일 이내에 각각 2페이지 정도의 심사보고가 작성되는데, 1개의 제안서당 대체로 4~5시간이 소요되며 이 심사보고가 CES(Center for Economics Studies)의 담당자에게 보내져 자료제공여부가 최종 결정된다. 주된 심사내용은 이용하고자 하는 데이터로 신청된 자료의 분석 가능 여부와 익명성 저축 여부이며 연구 내용이 센서스국의 직무내용에 일치할지에 대해서는 실질적으로 대부분 심사하지 않는다. usRDC의 이용이 허가된 후에는 SSS를 얻기 위한 법적인 수속으로 옮겨가고, 그 후 연방직원으로서의 고용계약이 체결된다. 최근에는 비밀보호의 문제가 중요시되어 이 수속에 많은 서류와 지문 등의 사적정보와 시간이 요구되며 이것들이 끝난 후에 각 usRDC에서 ID 카드가 작성되어 실제의 분석이 가능하

게 된다. CES에서는 이러한 절차를 고려하여 제안서의 제출로부터 데이터의 이용 가능까지의 기간을 6~12개월 정도로 어림잡아 연구계획을 세우도록 권장하고 있다.

### 3) 제공 후 관리

이용자는 usRDC의 데이터 서버에 접속하고 거기에서 모든 작업을 완결해야 한다. 각 단말기에 출력장치는 없으며 도출결과 등을 종이 혹은 전자매체로 usRDC 밖으로 반출할 경우 각 usRDC에 체류하고 있는 센서스국 직원의 승인을 거쳐야 한다. 분석 프로그램 등을 업로드할 경우도 같다. usRDC로의 출입기록과 데이터 서버의 로그는 모두 보존되어 누가, 언제, 어떻게 데이터에 접근하고 어떤 작업을 했는지를 사후적으로 확인하는 것이 용이하다. 작업 중에 제안서에 제안한 분석 내용을 변경할 필요가 생겼을 경우에는 원칙적으로는 변경 내용과 그 필요성을 제출하고 승인을 받을 필요가 있다. 그러나 이용하는 데이터를 변경할 경우나 근본적으로 다른 연구를 하는 경우 이외는, 사실상 사후적인 승인절차를 거치고 있다. 제공규정을 위반했을 시에는 강한 법적 제재를 받게 되는데 비밀보호에 대한 위반시 징역 5년 또는 250,000달러 이하의 벌금(혹은 그 양쪽을 병과)이 부과된다.

또한 단순한 분석결과뿐만 아니라 데이터를 이용해 나오는 최종 성과물에 대해서도 엄격히 관리하고 있는데 최종 성과물은 비밀보호에 관한 심사위원회에서 승인을 얻은 후, CES의 discussion paper로 공표되며, 학술지에의 투고가 허가된다. 대개의 경우 회귀 분석 및 각 변수의 1차·2차의 요약 통계량이 비밀보호에 저촉될 경우는 드물다. 그러나 다중크로스 표나 산포도를 작성했을 경우는 비밀보호에 저촉될 수 있는 가능성이 많으므로 심사가 엄중하고 따라서 성과물을 공표할 때까지의 시간이 많이 소비된다. 요약 통계량이라 할지라도 중간값이나 최대값, 최소값 등 특정 샘플을 내보일 경우에는 심사의 대상이 된다. 그래서 CES는 논문으로 게재하는 크로스 표는 될 수 있는 한 공표 데이터를 이용하는 것을 권장하고 있다.

## 2. 미국 노동통계국(Bureau of Labor Statistics: BLS)

미국의 고용·물가 등의 경제통계를 담당하는 연방기관인 노동통계국에서도 마이크로데이터를 제공하고 있는데, 어느 통계가 어느 방법으로 공개될지는 자료공개검토위원회(Disclosure Review Board)에서 조사항목마다 결정한다. 동일한 조사 내에서 공공이용파일(Public Use File)로 제공되는 부분, Off-site Research<sup>7)</sup>에서의 사용이 인정되는 부분, On-site에서만 인정되는 부분, 전혀 사용을 인정받지 못하는 부분이 혼재하는 경우도 자주 있으며, 조사시에는 어느 방법으로 어느 항목이 공개될지가 사전에 피조사자에게 설명된다. 사후적으로 공개방법을 변경하는 것은 어렵지만 자료공개검토위원회(Disclosure Review Board)의 결정에 근거한 변경은 인정하고 있다. 자료공개검토위원회는 내부조직이며 외부자는 관여하지 않는다. 단, 외부전문가가 관여하는 자료공개기술자문위원회(Technical Review Board)가 별도 설치되고 있어, 외부자가 통계조사의 공개 방법에 관해서 의견을 진술하는 기회가 확보되어 있다.

### 가. 공공이용파일(Public Use File)

공공이용파일로 제공되는 자료는 가계조사인 Consumer Expenditure Survey와 노동력조사인 CPS(Current Population Survey), 패널데이터인 NLS(National Longitudinal Survey)가 있으며, CPS와 NLS는 센서스국과 공동으로 제공되는 자료이다. 이 자료들은 노동통계국(Bureau of Labor Statistics: BLS)에서 전자매체로 구입이 가능하며, 앞의 센서스국의 Public Use File에서 소개한 바와 같이 ICPSR(Inter-University Consortium for Political & Social Research) 등 각지의 데이터 아카이브에도 기탁되고 있어 교육이나 연구에 빈번하게 이용되고 있다.

7) 자료제공 형식 중 Licensed File에 해당하는 경우로 특정한 심층이용자에게 제공되며 제공 전에 법적인 조치(서명 등)가 필요한 형식이다.

## 나. Off-site Research(Licensed File)

인가파일(Licensed File) 형식으로 제공하는 Off-site Research는 보다 엄격하게 보호할 필요가 있는 조사항목에 대해서 일정 절차를 거쳐 특정 개인에게 데이터를 제공하는 방법으로, BLS에서는 엄격한 절차를 두고 이 방법으로 데이터를 제공하고 있다.

BLS가 관할하는 통계 안에서 이 방법으로 이용 가능한 데이터는 National Longitudinal Survey Geocode Data와 Census of Fatal Occupational Injuries 등이 있는데, National Longitudinal Survey Geocode Data는 NLS79(National Longitudinal Survey of Youth 1979), NLS(National Longitudinal Survey) Young Adult, NLS97(National Longitudinal Survey of Youth 1997)의 지리적 항목을 중심으로 한 조사항목을 말한다. 이 조사항목은 각 조사에 있어서의 응답자의 거주지정보 및 해당지역의 경제변수, 태어난 시점의 거주지정보, 14세 시점의 거주지 정보, 이혼 후 양육에 관한 합의가 성립한 거주지 정보, 출생·결혼·이혼·입학 등의 연월일, 입학대학의 명칭과 소재지, 박사학위 취득시 전공명, 양친의 거주지에서의 거리 등이 있다.

자료의 이용 자격 요건으로는 미국 시민일 필요는 없지만 미국 주재로 소속기관이 미국에 있어야 하며, 이용자와 BLS 사이에는 BLS Agency Agreement가 체결되며<sup>8)</sup> 동시에 이용자의 소속기관과 BLS에도 합의 계약이 행해진다. 소속기관에는 기관장(대학이면 부학장 이상)의 승인을 얻는 것 이외에 Project Coordinator라고 불리는 실질적 책임자를 지정한 뒤 서버 등 전자환경, 파일이나 CD-ROM 보존을 위한 물리적인 준비를 갖추는 것이 요청된다. 소속기관과 BLS 사이의 계약서에는 계약 위반이 있었을 경우에는 소속기관이 민사상의 손해배상 의무를 지는 것도 명기되어 있다. 따라서 비밀보호와 관련한 위반 등이 있었을 경우, 이용자가 연방정부의 Agent로서 형사·민사소추를 받을 가능성이 생기는 것은 물론 이용자의 소속기관도 민사소추를 받을 수 있게 된다.

이용 기간은 학생이 1년이며 연구자는 3년 또는 소속기관에 재직 중인 기간이다. 이용 장소는 소속기관 내의 BLS가 지정한 장소이며 이용

---

8) BLS의 무급 Agent가 되는 것의 합의 계약.

자의 자택 등은 인정받지 못한다. 이용자가 소속기관을 변경했을 경우 이용 계약은 그 시점으로 끝난다.

제공 후의 관리도 상당히 엄격한데 자료 이용 종료시에는 소속기관 책임자에 의한 데이터 CD의 반환과 데이터 소거의 확인, BLS에 의한 성과물의 검열이 요청된다. 또한 이용 기간 동안에 BLS 직원에 의한 현장사찰이 행해지며 사찰시에는 특히 데이터 보존의 물리적 환경, 서버나 main frame의 보안 환경 등이 체크된다. 현재까지 Off-site Research에 있어서의 심각한 계약 위반은 데이터를 자택에 가져간 경우가 1건 있었다. 이 경우 이용 계약이 중단되어 해당 이용자, 해당 소속기관의 재이용이 금지되었으며, 이용자에 대한 형사소추, 소속기관에 대한 손해배상 청구까지는 이르지 않았다. 사찰로 밝혀진 위반의 대다수는 CD-ROM을 보관하고 있는 캐비닛 열쇠를 잘 사용하지 않은 경우와 데이터를 네트워크 상 공유해 다루는 등이었으며, BLS 직원이 그 자리에서 개선을 요구하고 소속기관이 대응·보고하는 것으로 수습되었다.

#### 다. On-site(Data Laboratories)

BLS의 On-site 이용 조건은 첫째 미국 시민일 것, 둘째 자격기관에 학생으로 재적하고 있거나 또는 고용되어 있는 것인데, 첫 번째 조건은 그다지 엄격히 해석되지 않고 두 번째 조건이 엄격하게 해석되고 있다. 두 번째 조건의 자격기관은 미국고등교육기관, NPO(국세청(IRS)이 정하는 기준 501c (1) 또는 501c (3)에 적합한 조직), 지방정부 또는 지방정부의 부속 기관, 지방정부를 대표로 하는 기관, 연방정부 연구기관 등이다. 외국국적의 객원교수라도 체류 연구기관이 자격기관에 포함될 경우에 한하여 이용 자격 인정을 받는다. 제공되는 자료의 목록은 부록으로 첨부한다.

제안서 제출에서 자료이용까지의 절차를 살펴보면, 먼저 5~10페이지 분량의 제안서를 제출하고 제안서에는 다만 통계적인 이용을 목적으로 하는 것, 연구 내용에 대해 통계 기술적인 메리트가 있는 것, BLS의 흥미가 있는 것, BLS의 정책을 진척시키는 것을 나타낼 필요가 있다. 그 밖에 최종 성과물의 공개 방법, 보조금, 공동 연구자의 유무 등을 써 넣

는 것이 요구된다. 제안서의 제출 마감은 연 4회(1월, 4월, 7월, 10월의 각 15일)이며 심사에 2개월, 심사 후 법적 계약 등에 2개월을 필요로 한 경우가 대부분이다. 제안서를 접수하는 시기는 자료이용까지 3개월부터 6개월을 계획하도록 한다.

제안서에 대한 심사는 MARB(BLS-wide Microdata Access Review Board)에서 하는데, BLS 내에서 데이터의 보안을 담당하는 직원, 이외 피신청 데이터를 관할하는 그룹의 시니어·이코노미스트가 출석하여 심사를 담당한다. 이용기간은 신청시 최장 2년간이며 한 번 더 연장 가능하여 최장 4년 동안 이용 가능하다. 이용시간은 월요일부터 금요일까지의 8시부터 17시까지이며, 별도 신청하면 아침 6시부터 이용할 수 있다.

BLS On-Site 서버에는 SAS 및 STATA가 인스톨되어 있으며 랩탑컴퓨터(laptop computer) 등의 반입은 인정하나 인터넷이나 단말기에의 접속 단자는 설치되어 있지 않고 가져온 전자기기를 어떠한 방법으로도 인터넷에 접속하는 것은 금지되어 있다. 단말기에 허용되는 외부출력은 화면만으로 한정되며 종이매체·전자매체에서의 출력은 할 때마다 BLS 담당자의 검열을 받을 필요가 있다. On-site의 서버가 설치된 장소의 출입에는 미리 등록된 ID 카드를 사용한 인증이 필요하며, 이것들을 통해 언제, 누가, 어디에서, 어느 데이터에 접근하여 무엇을 했는지를 알 수 있다.

BLS의 On-site는 1997년 9월 이후 현재의 방법으로 정식으로 운영된 이래 총 120명의 연구자가 이 제도를 이용하고 있으며, 2006년 기준 연간 이용 중인 연구제안서는 36개이며 연구자는 50명이다.

### 3. 캐나다

캐나다는 통계에 있어 집중형국가로 행정정보에 접근하는 권한도 통계법에 의해 주어져 있는 등, 정부통계에 관한 강력한 권한을 통계청이 가지고 있다. 그에 따라 공개 정보의 비밀보호에 관한 요청도 강해서 이와 관련한 위반에 대해서는 형사죄가 성립하는 등 강한 법적구속력을 가진다. 권한 없는 정보의 접근을 시도할 경우 6개월 이하의 징역 또는

1,000달러 이하의 벌금, 혹은 그 양쪽을 병과하며, 자산 가격에 영향을 미치게 하거나 투기행위를 행하거나 하는 등 개인의 특정 정보를 경제 거래에 이용하기로 했을 경우는 5년 이하의 징역 또는 5,000달러 이하의 벌금, 혹은 그 양쪽을 병과한다.

### 가. 공공이용파일(Public Use File)

각 조사의 공공이용파일 작성시에 마이크로데이터 제공 위원회(Microdata Release Committee)가 제공하는 통계의 범위나 조사항목을 검토한다. 제공되는 데이터로는 센서스 3% 샘플 자료, 가계소비조사자료 등이다. 2002년 4월 2일부터 원격접속파일(Remote Data Access) 방법으로도 제공하고 있으며, 비용은 2001년 센서스 자료가 1,500달러로 개인 파일은 1,000달러이고 가족파일은 300달러, 가구 및 주택파일은 200달러이다. 2002년 가계소비조사의 경우는 3,210달러이다. 자료 이용 전에 마이크로데이터 허가 협정(Microdata Licence Agreement)이 필요하다.

통계연구목적을 위해 개별정보들을 연결한 레코드 연결 데이터(Record Linkage Data)도 일부 제공하고 있는데, 이 경우 일반 Public Use File보다 정보공개를 보다 신중하게 처리한다. 1986년부터 레코드 연결에 관하여는 Agency's Policy on Record Linkage가 담당하고 있으며 이 자료의 이용 조건은 단지 통계적인 목적으로 사용하여야 하며, 자료 이용으로 인하여 창출되는 공공의 이익이 개인정보보호의 이익보다 크다고 판단되는 경우이어야 하는데 이는 검토위원회(Review Committee)의 승인 후 통계청장(Chief Statistician)의 승인 및 장관(Minister)의 승인을 얻어야 한다. 통계청장의 승인을 거칠 때는 비밀보호위원회(Office of the Privacy Commissioner)와 공공이익단체(public interest group)의 자문을 구하여 정한다.

캐나다에서는 특징적으로 DLI(Data Liberation Initiative)를 운영하고 있는데, DLI에 참여하는 캐나다 교육기관(대학, 연구소 등)이 연간비용을 부담하고 학생과 임직원들이 무제한으로 마이크로데이터를 사용하도록 하는 제도(대부분 도서관과 정보센터에 위치)를 두고 있다. 이는 1980년대에 마이크로데이터 비용이 높아져 연구자들이 상대적으로 가격이 저렴한 미국, 영국, 중국 자료를 이용하자, 캐나다의 현실에 대한 연구가 이루어지지 않는 문제가 발생하여 1989년에 센서스 가격의 상승

을 막고자 공공 데이터 이용자 연합(Canadian Association of Public Data Users)과 연구 도서관 연합(Canadian Association of Research Libraries)에 의해 임시구매컨소시엄(ad hoc buying consortium)이 조직되고, 1993년에 사회과학연맹(Social Science Federation of Canada)의 지원으로 작업반(Working Group)이 구축되어 DLI프로그램을 제안하여 1996년에 재무부 이사회(Treasury Board)의 승인을 거쳐 DLI 5개년 시험 프로젝트가 성공, 2001년 3월 DLI 프로그램이 영구적으로 정착되었다.

또한 DLI 외부자문위원회(DLI-EAC, External Advisory Committee)가 운영되고 있는데 이는 교육기관, 도서관 대표로 구성된다. 조직구성은 캐나다 통계청 외부에서 9명의 투표권자를 선정(지역 대표도서관, 학계, 연구자 대표)하고 캐나다 통계청 내부에서 5명의 투표권자를 선정한다(DLI 담당과장, 자료제공담당 등). 임기는 6년으로 1/3이 2년마다 바뀌며 66개 기관이 참가, 하부위원회를 둔다. DLI 외부자문위원회는 각 교육 기관 및 연구기관의 요구사항을 대표로 캐나다 통계청에 전달하며 마이크로데이터의 제공방법을 제안하거나 새로운 상품 서비스 형식을 제안한다. 또한 마이크로데이터와 관련한 연간보고서를 매년 6월 30일까지 작성하여 캐나다 통계청장에게 제출 발표하며 이 자료는 DLI 웹사이트에 게시된다.

#### 나. On-site(Data Laboratories)

캐나다의 On-site 방법은 미국의 방법을 벤치마킹하여 전반적인 부분에서 거의 비슷하게 운영된다. caRDC(Canada Research Data Center)라고 부르며 이용시 데이터 이용자는 선서를 하고 캐나다 통계청과 무급의 고용계약을 체결한다.

2007년 1월 기준, 캐나다 국내에 본소와 출장소를 합해 20개를 운영하고 있으며 이들 시설은 법적으로 캐나다 통계청의 건물로 차용, 각 시설에는 캐나다 통계청 직원이 적어도 1명은 상주하고 있다. 이용자 개인의 PC 지참, PDA나 휴대폰 등의 사용은 원칙적으로 허용되지 않는다. usRDC가 연구목적의 이용에 거의 한정되는 반면, caRDC는 주로 대학원 수준에서의 교육목적 이용도 적극적으로 추진하고 있다.

이용자격 요건으로는 첫째 정부기관 또는 RDC 네트워크 참가기관에 소속되어 있을 것, 둘째 연구자(소속이 있으면 국적은 묻지 않는다), 연방정부기관의 피고용자(주지방정부피고용자는 별도상담), 대학원생의 석사논문 또는 박사논문 집필에 일부 이용하는 자, 대학원생의 석사논문 또는 박사논문과는 관련되지 않고 이용하는 자, 학부학생의 최종학년에 속하는 자 중 어느 하나에 포함될 것이다. usRDC와 비교하여 대학원생의 이용이 특히 명확히 규정화되어 있는 점과 학부학생에게도 제공하고 있는 점이 특징이다. 소속이나 국적이 캐나다와 관련되지 않는 연구자도 전혀 이용할 수 없는 것은 아니고 별도 상담으로 결정되고 있어, 상당히 광범위하게 이용자를 상정하고 있다고 말할 수 있다. BLS나 usRDC와 다르게 통계청 업무와의 관련을 기록할 필요는 없다.

이용 절차는 웹사이트 상으로 사용자 계정을 작성하고 연구계획서와 이용자 본인 및 공동 이용자의 이력서 등을 첨부하여 신청한다. 신청된 연구계획서는 선별 검토 위원회(Selection and Review Committee)에 회부되고 2명의 연구자와 캐나다 통계청 직원이 심사한다. 연구적인 부분과 비밀보호적인 부분을 나누어 심사하는 것이 특징인데, 심사담당 연구자는 관련 분야에서 뽑혀 연구목적과 데이터분석 수법의 합리성을 주로 심사하며, 통계청 직원은 On-site로의 접근 필요성이나 분석 결과가 비밀보호에 영향을 미치는 부분을 주로 심사한다. 제공되는 자료의 목록은 <부록 C>로 첨부한다.

제공 후의 관리 부분을 살펴보면, 연구 수행 중 반출하는 중간성과물에 대해서는 캐나다 통계청의 승인을 얻을 필요는 있지만 캐나다 통계청에 제출할 필요는 없다. 그러나 최종성과물은 제출해야 한다. 어떤 형태로 제출할 것인가는 계약시에 미리 지정하며 RDC working paper, Peer-reviewed journal article, Book or Book chapter, Thesis or Dissertation, Commissioned Report 중 선택한다.

2006년 776개의 프로젝트, 이용자 853명(516명의 연구자, 337명의 학생), 1,225개 데이터(NLSCY<sup>9)</sup>: 245개, CCHS<sup>10)</sup>: 223개, SLID<sup>11)</sup>: 143개의 순)가 신청되어 활용되고 있다.

9) National Longitudinal Survey of Children and Youth.

10) Canadian Community Health Survey.

11) Survey of Labor and Income Dynamics.

## 4. 호 주

특정 개인이나 기업에 대한 정보를 식별해 낼 수 없도록 비밀보호 방법<sup>12)</sup>을 적용한 Public Use File을 CURF(Confidentialised Unit Record File)라 부른다. CURF는 세분화 정도(levels of detail)에 따라서 3가지 종류로 분류되며 그에 따라 제공방법을 달리하고 있다.

### 1) 제공방법

CURF의 세분화 정도는 기본형(Basic), 확장형(Expanded), 전문형(Specialist)으로 분류되며 전문형이 가장 세분화된 자료이다. 기본형(Basic)은 CD-ROM, RADL(Remote Access Data Laboratory), ABSDL(Australian Bureau of Statistics Site Data Laboratory)로 접근 가능하며, 확장형(Expanded)은 RADL, ABSDL로 접근 가능하며, 전문형(Specialist)은 ABSDL로만 접근 가능하다. RADL은 허가 받은 이용자가 호주 통계청(Australian Bureau of Statistic: ABS)이 제공하는 마이크로데이터를 인터넷을 통해 제공받을 수 있는 방법을 말하며, 매년 사용자의 비밀번호가 변경되며 승인된 데이터에만 접근이 가능하다. ABSDL은 ABS의 감독 하에 ABS 내에서 마이크로데이터를 이용하는 방법으로, 향후 먼 지역의 사용자를 위해 ABS 외부 데이터실험실(Data Laboratories: DL)을 개발할 예정이다. 전문파일의 경우 신청하기 전에 호주 통계청(ABS)의 마이크로데이터 담당자와 접촉해야 하며 각각의 경우에 따라서 CURF의 제공 여부를 결정한다.

호주 통계청에서는 확실한 수요(demonstrated user demand)가 있을 때 생산하고 제공되고 있는 CURF 리스트를 홈페이지에 게시하며 향후 서비스될 CURF 목록과 일정을 게시한다. CURF 사용자 훈련 매뉴얼(Training Manual), 사용자 가이드(User Guide)를 제공하며, 마이크로데이터 관련 뉴스레터를 분기별로 발간하고 홈페이지 상의 FAQ도 활성화하는 등 마이크로데이터 제공 관련 홍보도 중요시하고 있다. 외국인에 자료 제공은 경우에 따라 다르지만 외국인이라고 해서 기본적인 자료만

---

12) 이름, 주소 등 제거, 값 제거(이상치, 작은 값 등), 셀 변경, 통합(Randomly, Swap, Drop, Substitute)기법 등 적용.

접근할 수 있는 것이 아니라 전문(Specialist)파일에도 접근이 가능하다. 호주 통계청에서 제공하는 전체 자료 목록은 <부록 D>를 참조하기 바란다.

<표 3-5> 호주의 제공방법별 이점 및 제약사항

	CD-ROM (기본형)	RADL (기본&확장형)	ABSDL (전문형)
데이터 세분화 정도	세분화 정도가 약함	다소 세분화됨	상당히 세분화됨
이점	<ul style="list-style-type: none"> <li>· 이용자가 자신의 컴퓨터에서 제약 없이 사용 가능</li> </ul>	<ul style="list-style-type: none"> <li>· 인터넷이 가능한 모든 컴퓨터에서 접근 가능</li> <li>· 좀 더 세분화된 데이터 분석이 가능</li> <li>· 시스템 내에서 SAS, Stata와 SPSS 소프트웨어가 제공됨</li> </ul>	<ul style="list-style-type: none"> <li>· ABS내 관리감독 하에 지정된 컴퓨터에서 접근 가능</li> <li>· 높은 수준의 데이터 분석이 가능</li> <li>· 시스템 내에서 SAS, Stata와 SPSS 소프트웨어가 제공됨</li> <li>· 시스템에서 다른 데이터 세트들과의 결합도 허용될 수 있음</li> </ul>
제약사항	<ul style="list-style-type: none"> <li>· 법적 서약서(legal undertaking)에 의해 구속됨</li> </ul>	<ul style="list-style-type: none"> <li>· 법적 서약서에 의해 구속됨</li> <li>· 개별 레코드는 ABS IT환경 내에 남아 있음</li> <li>· 쿼리의 속성과 결과물의 성격과 크기에 따라 제약이 있음</li> <li>· 이용자의 활동이 모니터 됨</li> </ul>	<ul style="list-style-type: none"> <li>· 법적 서약서에 의해 구속됨</li> <li>· 개별 레코드는 ABS IT환경 내에 남아 있음</li> <li>· 쿼리의 속성과 결과물의 성격과 크기에 따라 제약이 있음</li> <li>· 이용자의 활동이 모니터 됨</li> <li>· ABS 내에서 ABS의 감독을 받아야 함</li> </ul>

자료: 통계청 국제통계협력과, “외국인에 대한 마이크로데이터 제공검토”, 2004에서 인용.

## 2) 가격책정

CURF의 가격책정은 세분화 정도와 자료 제공방법에 따라 <표 3-6>과 같이 차등 책정되고 있으며, 이용자별 가격책정이 아닌 기관별 책정을 하고 있다. 호주 통계청 Microdata Access Strategies Section의 디렉터인 Kim Farley-Larmour에 따르면, 호주의 마이크로데이터 가격책정은 호주정부 가이드라인에 따라 원가보상원칙(cost recovery ABS Pricing Policy)을 기준으로 책정하고 있다고 한다. 한 기관에서 다수의 이용자가 해당 마이크로자료를 함께 사용한다고 해서 가격이 올라가는 것은 아니지만 이용자 수를 제한할 수는 있다.

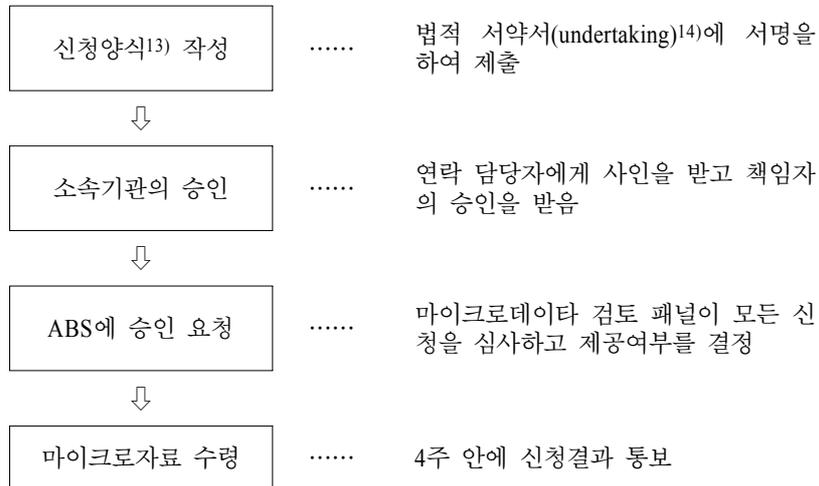
<표 3-6> 가격책정 기준

가격책정 기준	가격
기본형(CD-ROM)	\$1,320
기본형(RADL)	
확장형(RADL)	
기본형(CD-ROM/RADL)과 확장형(RADL) 묶음자료	\$1,980
컨설턴트(프로젝트당)	\$435
전문형(ABS DL)	용도에 따라 다름

자료: 호주통계청(ABS) 홈페이지.

## 3) 자료신청절차

사용자 개인, 연락 담당자(contact officer), 책임자(responsible officer), 컨설턴트, 그리고 컨설턴트 지원기관별로 구분하여 자료 신청절차 및 역할과 의무사항 등에 대한 일련의 과정들을 홈페이지를 통해 상세히 안내하고 있다. 대부분 승인된 기관에서 자료를 사용하고 있으며, 승인된 기관과 전혀 연관이 없는 개인의 CURF 접속은 불가능하다.



#### 4) 관계자별 역할과 의무

호주 통계청의 경우는 통계청에서 바로 개인으로 자료를 제공하는 것이 아니라 기관을 통해 제공하므로 관계자별 역할과 의무를 가진다. 관계자는 사용자 개인과 연락담당자, 책임자, 컨설턴트로 구분할 수 있는데, 먼저 사용자 개인의 경우는 신청서와 서약서를 작성하기 전에 반드시 CURF 훈련 매뉴얼(Training Manual)을 읽어야 하며, 신청서와 서약서는 소속기관의 연락 담당자의 서명을 받아 ABS에 제출해야 한다.

사용자의 의무는 그들이 서명한 서약서에 의해 통제되며, CURF는 반드시 명시된 통계적 목적으로만 사용해야 한다. 특정 개인이나 기관을 식별하려는 시도를 해서는 안 되며, 다른 사람이나 기관에 자료를 넘겨서는 안 된다. 승인된 기간이 만료된 후이거나 소속기관으로부터 접근권이 박탈된 후 소속기관을 그만두었을 시에는 자료 사용을 금지하고 있다.

연락 담당자는 ABS 담당자와의 연락을 담당하며 사용자 개인의 CURF에 대한 접근을 도와주는 역할을 한다. 또한 소속기관에 대한 책임자의 서약서(ROU<sup>15)</sup>)를 조정하고, 사용자의 서약서(IUU<sup>16)</sup>)와 신청서

13) 신청하는 CURF 개수에 따라 신청 양식이 다름

14) 개인이나 기관의 책임자별로 서약서 양식이 다름

(REQ) 및 컨설턴트 서약서(CON17)) 작성을 도와야 한다. 그리고 기관이 보유한 CURF 목록을 관리하며, 원본 CD-ROM를 안전하게 보관하고 복사본을 개인 사용자에게 제공하여야 한다. 기간 만료 후 원본 CD-ROM을 ABS에 반납하고, 개인에게 제공되었던 복사본을 파기하여야 한다. 사용자가 소속기관을 떠나게 되면 사용자의 모든 CURF에 대한 접근권이 박탈되고 CURF를 반납하여야 하며 떠나기 최소 10일 전에는 ABS에 알려야 한다.

책임자의 경우는, 마이크로자료 이용신청시 책임소재를 명확히 하기 위해 반드시 각 기관의 장 또는 허가된 개인을 책임자로 지정하여 접촉해야 하며, 일반적으로 대학의 부총장, 정부의 기관장, 사기업의 CEO가 책임자로 지정된다. 컨설턴트의 경우, 각 기관이 기관에서 보유하고 있는 CURF를 사용하는 연구 프로젝트를 수행하기 위해 외부로부터 컨설턴트를 고용하는 경우 컨설턴트용 신청서와 서약서에 서명하면 ABS에서 컨설턴트 접근권을 제공한다. 사용자와 동일한 의무를 부여받으며, 컨설팅(용역 사업)이 완료되면 모든 CURF 접근권이 박탈되고 CURF를 파기하여야 한다. 컨설팅(용역 사업)이 완료되면 고용된 기관으로 CURF와 관련된 모든 분석 결과물들을 넘겨야 하며 이에 대한 소유권은 고용된 기관에 있다.

호주 통계청에서는 3명으로 구성된 Data Access & Confidentiality Methodology Unit에서 CURF를 생산하며, 개인의 정보를 보호하기 위하여 CURF를 어떻게 생성하는 것이 효율적인가를 연구하고 있다. 또한 4명으로 구성된 CURF Management(sub-section)에서는 생성된 CURF를 이용자에게 제공하고, 관계자들을 총괄하며 전반적인 관리를 하는 역할을 한다.

## 5) 제공 후 관리

제공 후 CURF를 이용하여 진행된 연구 결과물들을 홈페이지에 공개하고 Auditing Program을 통해서 결과물을 검토하고 있다. Auditing Program

---

15) Responsible Officer Undertaking.

16) Individual User Undertaking.

17) Consultant Undertaking.

은 ABS의 코드와 결과물을 감사(Audit)하여 개별 정보를 노출시키지 않도록 하며, RADL로 이용시 매년 어떤 자료를 이용하였는지, 접근횟수는 얼마인지 등을 체크하여 문제가 발생하면 이용 중단시키는 역할을 하고, 또한 결과물을 체크하여 정보를 지나치게 많이 담은 경우에는 다시 ABS로 회수하고 향후 사용중단 조치를 취한다.

자료 제공에 관한 의무 위반시 CURF 접근권한을 박탈<sup>18)</sup>하고 이를 어길 시 13,200달러 이내의 벌금 또는 2년 이하의 징역에 처한다. CD-ROM의 수거는 잘 이루어지고 있으며 간혹 분실된 경우에는 차후 이용을 못 하게 한다.

호주에는 ABS-AVCC(Australian Vice Chancellors Committee) CURF 협약<sup>19)</sup>을 두어 학문적 연구와 대학수업을 지원하기 위하여 대학, 정부와 통계청이 협약을 맺어 CURF 이용을 활성화하고 있다. 해당기관의 연락 담당자(Contact Officer)는 CD-ROM 원본을 보유하고 사용자에게 복사본을 제공하고 수거하며 주로 교수, 사서, 관리자 등으로 지정된다. CD-ROM과 RADL에의 접근권이 포함되어 있으나 ABSDL 접근권은 포함되어 있지 않다.

## 5. 영 국

영국의 마이크로데이터 제공현황에 대해 각 기관별로 알아보고자 한다. 영국은 영국통계청(Office for National Statistics: ONS)의 주관하에 크게 4개의 기관 CCSR(Centre for Census and Survey Research), UKDA(UK Data Archive), CeLSIUS(Centre for Longitudinal Study Information and User Support), ONS-IDBR(Inter-Departmental Business Register)에서 마이크로데이터를 제공하고 있으며, 기관에 따라 주로 다루는 자료의 종류가 다르다. 각 기관에서 제공하는 자료는 모두 비밀보호기법이 적용된

18) 이전에 일부 위반사례가 발견되어 개인(대학 교수)으로부터 향후 규정을 준수하겠다는 서약을 받은 적이 있으나 중대한 위반사례는 아직 없었음.

19) 대학교에서 CURF를 상업적 목적으로 사용하려는 경우에는 이 협약이 적용되지 않음. 또한 이 협약을 맺지 않은 대학교는 현행 ABS의 CURF 가격을 지불하고 구매할 수 있으며, 학문적 연구 목적으로 CURF를 이용하여 그 결과물을 발간할 경우 이를 인정함.

마이크로데이터이며, 상세하고 개인 식별이 되는 자료는 ONS의 담당자와 연락하여 제공여부를 결정한다.

### 가. CCSR(Centre for Census and Survey Research)

ESDS(Economic and Social Data Service) Government의 주도적인 역할을 하는 CCSR은 1992년 4월에 설립되어 맨체스터(Manchester) 대학 내에 위치하고 있으며 주로 인구센서스 자료를 제공한다. 1991년, 2001년의 Census로부터 작성된 SAR<sup>20)</sup>을 개인과 세대의 2종류로 제공하고 있는데 제공되는 마이크로데이터로는 2% 개인 SAR(약 110만 레코드), 3% 개인 SAR(약 175만 레코드), 1% 세대 SAR(약 20만 레코드, 세대원수 약 50만 명), 5% SAM<sup>21)</sup>(약 300만 명의 개인 레코드)이며, 보다 상세한 개인 데이터인 CAM(Controlled Access Microdata Samples)은 ONS 내에 On-site로만 이용이 가능하다. On-site로 이용시 좀 더 까다로운 절차를 가지는데, 구체적으로 설명하면 개인 CAM은 CRAB(Census Research Access Board)의 심사를 받은 후, ONS의 마이크로데이터공개위원회(Microdata Release Panel: MRP), GROS(General Register Office for Scotland), NISRA(Northern Ireland Statistics and Research Agency)의 심의 후 이용 가능하다. 영국의 국가특성상 잉글랜드, 스코틀랜드, 북아일랜드 각각의 심의기구를 모두 거쳐야 한다. 외국인에 대해서는 3% 개인 SAR과 5% SAM 데이터에 한정해 이용을 승인하고 있다. 개인 CAM 자료를 제외한 1991 SAR, 2001 SAR, 2001 SAM의 이용시에 학술목적으로 이용하는 경우는 아이디와 패스워드를 취득한 다음 CRS(Census Registration Service)에 등록하여 이용 가능하며, 비학술목적으로 이용하는 경우는 End user Licence Agreement에 서명하여야 한다.

데이터 수집, 제공, 분석에 있어서의 비밀 보호에 관한 연구를 실시하는 CAPRI(Confidentiality and Privacy)라는 팀을 별도 운영하고 있는데, 현재 Dr. Kingsley Purdam이 주도하며 맨체스터 대학에 소속되어 있는 사람들로 구성되어 있다. 마이크로데이터나 소지역 집계 데이터의 리스

20) Sample of Anonymised Record: 비밀보호기법이 적용된 센서스 마이크로데이터를 SAR이라 부른다.

21) Small Area Microdata file.

크 레벨을 입증하는 혁신적인 방법을 개발해 ONS의 센서스나 표본조사의 리스크 평가에 막대한 공헌을 하였으며, 국제 심포지엄도 개최하는 등 활발히 활동하고 있다.

이용요금은 학술목적과 비학술목적으로 크게 나누어지는데, 학술목적의 경우 무료로 제공되며 CD-ROM으로 구입시에는 30파운드가 부과된다. 비학술목적의 경우에는 데이터 세트당 1,000파운드의 높은 가격이 책정되며, 중앙정부, 지방정부, 보건기관, 자선단체의 경우에는 500파운드가 부과된다. 각각의 경우에 17.5%의 부가가치세가 별도로 부과된다.

#### 나. UKDA(UK Data Archive)

1967년에 설립되어 에섹스(ESSEX) 대학 내에 위치한 UKDA는 주로 각종 사회, 인문에 관한 서베이 데이터를 제공하고 있다. 제공되는 자료는 사회·인문 전반의 각종 서베이 데이터 및 2001 세대 SAR(Sample of Anonymised Record)이다. ARC(Acquisitions Review Committee)가 격주마다 개최되어 데이터 제안서를 심사한다.

마이크로데이터의 이용요금은 학술목적인 경우 무료이며, 비학술목적은 500파운드가 부과된다. CD-ROM으로 요청할 경우 기본료 7.5파운드에 수수료(handling fee) 2.5파운드가 부과되며, 우송비가 따로 추가되는데 영국 내에서는 3파운드, 영국 이외 유럽은 4파운드, 그 외 지역은 5파운드이다. 역시 17.5%의 부가가치세가 부과되며, 교육목적으로 사용하는 연구생과 학생, EU의 이용자, 각 학부 내의 목적(departmental purpose)으로 이용하는 에섹스 대학의 대학 관계자는 세금이 면제된다.

#### 다. CeLSIUS(Centre for Longitudinal Study Information and User Support)

CeLSIUS는 2001년 12월에 설립되어 런던시립대학(London City University) 내에 설치되어 있으며 LS데이터(Longitudinal Study Data)를 제공하고 있다. LS데이터는 1974년부터 작업을 시작하였으며 직업별 사망데이터의 타당성이나 장기적 출생데이터의 필요를 시작으로 작성되었다. 이 데이터는 링크지된 1%(약 50만)의 잉글랜드와 웨일스의 센서

스개표데이터이다. 특정 4개의 탄생 월에 태어난 사람을 LS샘플로 지정하여 센서스정보와 다른 등록 정보를 연결하고 사망, 출산, 이주, 암환자 등의 다른 정보들을 추가하여 패널 데이터 형식으로 구축한 자료이다. 잉글랜드와 웨일스의 데이터를 제공하고 있으며 스코틀랜드는 현재 작성 중이다.

자료의 이용에 관해서는 LSRB(Longitudinal Study Research Board)에서 심의한다. 영국내 학술기관 이용자 이외에는 ONS의 LS담당자와 직접 접촉하여야 한다. 자료이용 신청시 연구의 논점이 명확해야 하며, CeLSIUS의 직원과 연구과제 및 연구기간에 대해 상담하고 온라인 트레이닝을 실시한다. 비밀보호(Confidentiality)에 관한 계약서(LS License Agreement)에 서명을 한 후 연구계획서(form A)를 제출하고 CeLSIUS 직원의 지시에 따라 Data Dictionary를 이용해 변수를 추출하고 추출된 변수를 바탕으로 신청서(form B)를 작성해 ONS LSRB에 제출한다. ONS LSRB의 이용 승인 후 계약서를 제출하고 가공 데이터를 수령한다. 데이터와 관련한 자료를 발표할 때는 사전 final outputs clearance를 취득해야 한다.

이용요금은 영국내 학술기관의 교직원 및 학생은 무료이며 그 외에는 LSRB에서 결정한다.

## 라. ONS-IDBR(Inter-Departmental Business Register)

ONS 내부의 ONS-IDBR에서는 사업체자료(Business Data)를 제공하고 있다. 제공되는 데이터는 기업 활동 분석의 주된 자료로 이용되고 있으며 기업을 대상으로 하는 표본조사의 표본추출틀(sample frame)도 제공하고 있다. 자료의 제공은 MRP(Microdata Release Panel)를 운영하여 심의를 거쳐 제공한다. 중앙정부의 직원 및 중앙정부와 계약한 사람은 조사 및 분석을 위한 목적으로는 개인식별가능 자료(Disclosive Data)의 이용이 가능하며, 지방정부 및 다른 정부단체의 직원은 일부의 개인식별가능 자료(Disclosive Data) 이용이 가능하고, 일반인의 경우는 개인식별불가능 자료(Non-disclosive Data)의 이용이 가능하다.

이용요금은 개인식별가능 자료와 개인식별불가능 자료, 그리고 개인

식별가능 자료의 표본자료를 사용할 경우에 1시간당 60파운드가 부과되며, 개인식별가능 자료의 구역자료를 사용할 경우는 구역당 150파운드가 부과된다.

## 6. 일 본

일본의 마이크로데이터 제공은 정부기관에 대해서만 지극히 제한적으로 이루어졌다. 이에 따라 마이크로데이터 제공에 관한 시스템이 다소 미약하다. 최근 마이크로데이터에 대한 요청이 증가함에 따라 2009년 제공을 목표로 ‘통계데이터의 2차 이용 촉진에 관한 연구회’를 개최하는 등 마이크로데이터 제공 서비스를 준비하는 데 많은 노력을 하고 있다. 현재 히토쓰바시대학 경제연구소 부속 사회과학통계정보연구센터에서 정부통계 마이크로데이터 제공 시범시스템을 구축하여 운영하고 있는데 이를 간략히 소개하는 것으로 일본의 사례를 대신한다.

2004년 11월에 시작하여 2008년 9월까지의 계획으로 마이크로데이터 제공에 관해 시범 운영하고 있다. 이용 장소는 히토쓰바시대학 사회과학통계정보연구센터이며 이용조건은 학술연구에 한정하고 교육목적 및 영리기업의 위탁연구는 허가하지 않는다. 신청자는 전국 대학의 강사 이상의 전임교원으로 하며 연 4회(1월, 4월, 7월, 9월) 이용자를 모집한다. 이용자는 신청서를 다운로드 하고 센터에 신청한다. 심사 후 센터에서 마이크로데이터를 작성하고 설명회를 개최한 후 이용자는 데이터를 이용하게 되며, 이용기간 종료 후 PC 내의 데이터를 소거하고 CD는 반환한다. 데이터의 사용기간은 6개월이며 이용자는 마이크로데이터를 이용해 작성한 연구논문을 센터에 2부 제출해야 한다. 2007년 7월 기준 현재 이용건수는 66건, 이용자 수는 92명이다.

## 제4절 시사점

앞에서 우리나라 통계청의 제공현황과 해외 여러 나라의 마이크로데이터의 제공현황에 대해 살펴보았다. 이를 바탕으로 다음과 같은 시사

점을 도출하였다.

첫째, 해외의 마이크로데이터 제공방법에 있어 공공이용파일(Public Use File)로 제공 가능한 자료는 간단한 절차로 제공하고 있으며, 특별한 학술연구를 목적으로 하는 자료는 On-site방법을 통해 제공하는 것이 추세로 보인다. 또한 ‘기관 대 개인’이 아닌 ‘기관 대 기관’의 형태로 제공되고 있음을 눈여겨볼 필요가 있다. 기관을 상대로 하여 제공하는 형태는 개인이 자료 이용시 비용의 부담이 적을 뿐만 아니라 자료의 제공을 담당하는 기관에서도 관리적인 측면에서 더 효율적일 것이라 판단된다.

둘째, 제공절차 면에서 On-site방법으로 자료를 제공할 경우에는 제안서(proposal)의 심사뿐 아니라 최종 성과물 관리까지 엄격하게 관리하고 있는데, 우리나라에서도 이용 중의 상황이나 최종성과물에 대한 보다 철저한 관리감독 체계가 요구된다. 또한 자료제공위원회를 운영함에 있어 통계작성자의 입장과 이용자의 입장을 원활하게 조율하기 위해 외부전문가를 통한 자료제공기준이나 제공방법, 그리고 새로운 상품이나 서비스 방식 등에 대한 의견수렴의 체계적인 절차가 필요하다.

셋째, 마이크로데이터 제공비용 산정시 이용목적별 또는 기관단위별로 가격을 차등화할 필요가 있다. 연구, 자금지원연구, 발간, 소득창출 등의 각각의 경우에 따라 가격을 다르게 산정하거나, 대학이나 연구소 등의 기관단위로 일정기간을 정해 비용을 산정하는 방식의 검토가 필요하다.

넷째, 각국에서는 다양한 마이크로데이터의 제공과 그에 따른 비밀보호방법의 개발에 많은 노력을 기울이는 점에도 주목할 필요가 있다. 이용자의 데이터 이용 목적 및 수요를 파악하여 그에 따른 다양한 상품을 개발해야 하며, 응답자의 개인정보 보호와 응답거부 방지, 응답자와 통계작성기관 간의 신뢰관계 유지를 위해 제공되는 마이크로데이터에 대해 비밀보호방법을 확대 적용하여야 할 것이다. 한편 이용자를 고려하여 자료 정보의 손실을 최소화하는 비밀보호방법을 연구하여 적용해야 할 것이다.

다섯째, 홍보적 측면으로 보면 이용자의 편의를 위해 마이크로데이터 신청방법, 자료제공규정, 이용절차 등을 수록한 상세한 지침 안내서 등의 마련과 제공에 노력을 기울일 필요가 있을 것이다.

이상에서 외국사례를 통하여 제공방법, 제공절차, 비용 산정, 사후관리, 이용자 홍보에 대한 시사점을 도출하였으나, 무엇보다도 자료의 최대 이용자인 연구자들과 긴밀한 협조체제를 구축하여 단순한 데이터 제공자와 이용자의 관계에서 더 나아가 상호 발전할 수 있는 방향을 모색해야 할 것이다.

효율적인 마이크로데이터의 제공은 연구자들로 하여금 좀 더 복잡한 경제·사회문제들에 대하여 깊이 있는 연구를 가능하게 하며, 연구자들로부터의 피드백을 통하여 자료의 품질향상을 꾀할 수 있고, 이를 통하여 국가통계의 신뢰성을 구축할 수 있게 할 것이라 기대한다.

## 참고문헌

- 박원환·황조연(2004), 「통계자료의 특성과 비밀보호 방법에 관한 연구」, 통계청.
- 김창환(2007), 「IPUMS-Integrated Public Use Microdata Series」, Minnesota Population Center.
- 통계청, “통계자료제공규정”.
- 통계청 국제통계협력과(2004), “외국인에 대한 마이크로데이터 제공검토”.
- 통계청 통계정보국(2006), “통계 마이크로데이터 제공현황 및 개선방안”.  
\_\_\_\_\_(2006), “마이크로데이터 서비스 업무편람”.
- \_\_\_\_\_(2006), “마이크로데이터 활용도 제고를 위한 심층의견조사결과 보고서”.
- \_\_\_\_\_(2006), “OECD 마이크로데이터 관련회의 참가결과 보고”.
- \_\_\_\_\_(2007), “ON-SITE ACCESS 운영지침”.
- \_\_\_\_\_(2007), “On-Site Access 자료유출 방지 계획(안)”.
- \_\_\_\_\_(2007), “마이크로데이터 제공비용 산정기준 개정방안”.
- 神林 龍(2007), 「北米における政府統計個票調査公開の現状に関する調査報告 : 米國勞働統計局, 米國センサス局およびカナダ統計局のオンサイトリサーチを中心に」, Institute of Economic Research Hitotsubashi University.
- 統計データの二次利用促進に関する研究會(2007), “제1회 회의보고서”.
- \_\_\_\_\_(2007), “諸外國の統計データの二次的利用の狀況”.
- ABS(2005), “Responsible Access to ABS CURFs Training Manual”.
- \_\_\_\_\_(2006), “Remote Access Data Laboratory User Guide”.
- \_\_\_\_\_(2007), “Managing ABS Confidential Unit Record File(CURFs) A Step By Step Guide”.

London School of Economics(2003), “Providing remote access to data : the academic perspective”.

ONS(2004), “ONS Micro-data Access and Release Policy”, Statistical Policy Committee.

\_\_\_\_(2005), “Providing access to data and making microdata safe, experiences of the ONS”.

\_\_\_\_(2005), “Access to business microdata in the UK : dealing with the irreducible risks”.

\_\_\_\_(2007), “MRP Guide - ONS MICRODATA RELEASE PROCEDURE”.

<http://www.nso.go.kr/shop/>

<http://mdss.nso.go.kr/>

<http://www.census.gov/>

<http://www.bls.gov/>

<http://www.thedataweb.org/>

<http://ces.census.gov/>

<http://factfinder.census.gov/>

<http://dataferrett.census.gov/>

<http://www.ipums.org/>

<http://www.nber.org/data/>

<http://www.ceprdata.org/>

<http://www.unicon.com/>

<http://www.abs.gov.au/>

<http://www.statistics.gov.uk/>

<http://www.data-archive.ac.uk/>

<http://www.ccsr.ac.uk/>

<http://www.celsius.lshtm.ac.uk/>

## < 부 록 >

### A. 미국 usRDC 제공자료 목록(2006년 10월 기준)

1	American Community Survey
2	American Housing Survey
3	Annual Capital Expenditures Survey
4	Annual Survey of Manufactures
5	Auxiliary Establishment - ES9200
6	Business Expenditures Survey
7	Census of Construction Industries
8	Census of Finance, Insurance, and Real Estate
9	Census of Manufactures
10	Census of Mining
11	Census of Retail Trade
12	Census of Services
13	Census of Transportation, Communications, and Utilities
14	Census of Wholesale Trade
15	Commodity Flow Survey
16	Compustat-SSEL Bridge
17	CPS Crosswalk
18	Current Population Survey March Supplement
19	Decennial Census Long Form Sample
20	Decennial Employer-Employee Database
21	Economic Census of Puerto Rico
22	Enterprise Summary Report - ES9100 (large company)
23	LEHD Business Register Bridge
24	LEHD Employer Characteristics File
25	LEHD Employer Quarterly Workforce Indicators
26	Longitudinal Business Database
27	Manufacturing Energy Consumption Survey
28	Medicaid Eligibility Database System - California
29	Medical Expenditure Panel Survey - Insurance Component
30	National Employer Survey

(계속)

31	National Longitudinal Survey
32	Ownership Change Database
33	Quarterly Financial Report
34	SIPP Public-Use Crosswalk
35	Standard Statistical Establishment List - non Name and Address File
36	Standard Statistical Establishment List - Name & Address File
37	Survey of Income and Program Participation (SIPP) Panels
38	Survey of Income and Program Participation - Longitudinal
39	Survey of Industrial Research and Development
40	Survey of Manufacturing Technology
41	Survey of Plant Capacity Utilization
42	Survey of Pollution Abatement Costs and Expenditures
43	Unemployment Insurance-Base Wage File
44	Insurance Programs
45	Occupational Employment Statistics Survey
46	Current Employment Statistics
47	Job Openings and Labor Turnover Survey
48	Mass Layoff Statistics
49	National Compensation Survey and its predecessors, including:
50	Employment Cost Index
51	Employer Costs for Employee Compensation
52	Occupational Compensation Survey
53	Industry Wage Surveys
54	Area Wage Surveys
55	Employee Benefits Surveys
56	Survey of Occupational Injuries and Illnesses
57	Consumer Expenditure Surveys
58	Producer Price Indexes
59	Consumer Price Indexes
60	International Price Program
61	Quarterly Census of Employment and Wages Program, and the associated Longitudinal Database of Establishments covered by State Unemployment

## B. BLS On-site 제공자료 목록(2007년 9월 기준)

---

### Data on employment and unemployment :

- National Longitudinal Survey of Youth 1979 (NLS79) Zip Code and Census Tract Files
- National Longitudinal Survey of Youth 1997 (NLS97) Zip Code and Census Tract File
- NLS79 School Surveys
- NLS97 School Surveys
- National Longitudinal Surveys Original Cohorts: Mature and Young Women, and Older and Young Men files with Geographic Variables (2)
- 1993 and 1995 Surveys of Employer Provided Training
- Quarterly Census of Employment and Wages Program, and the associated Longitudinal Database of Establishments covered by State Unemployment Insurance Programs (3)
- Occupational Employment Statistics Survey
- Current Employment Statistics
- Job Openings and Labor Turnover Survey
- Mass Layoff Statistics

### Data on prices and living conditions :

- Consumer Expenditure Surveys
- Producer Price Indexes
- Consumer Price Indexes
- International Price Program

### Data on compensation and working conditions :

- National Compensation Survey and its predecessors, including:
    - Employment Cost Index
    - Employer Costs for Employee Compensation
    - Occupational Compensation Survey
    - Industry Wage Surveys
    - Area Wage Surveys
    - Employee Benefits Surveys
  - Survey of Occupational Injuries and Illnesses
-

## C. 캐나다 caRDC 제공자료 목록(2007년 1월 기준)

1	Aboriginal Peoples Survey (APS)
2	Canadian Community Health Survey (CCHS)
3	Canada Survey of Giving, Volunteering, and Participating (CSGVP)
4	Ethnic Diversity Survey (EDS)
5	General Social Survey (GSS)
6	Health Services Access Survey (HSAS)
7	Labour Force Survey (LFS)
8	Longitudinal Survey of Immigrants to Canada (LSIC)
9	National Longitudinal Survey of Children and Youth (NLSCY)
10	National Population Health Survey (NPHS)
11	National Graduates Survey (NGS)
12	Participation and Activity Limitations Survey (PALS)
13	Public Service Employee Survey (PSES)
14	Survey of Approaches to Educational Planning (SAEP)
15	Survey of Labour and Income Dynamics (SLID)
16	Workplace and Employee Survey (WES)
17	Youth in Transition Survey (YITS)

D. 호주 통계청 제공자료 목록(2007년 9월 기준)

조사명	기준연도	세분화 정도
Aspects of Literacy	1996	기본형*
Australian Housing Survey	1994, 1999	기본형*
Australians' Employment and Unemployment Patterns	1994-97	기본형*, 확장형
Business Longitudinal Survey	1994-98	기본형*
Census - Household Sample File	1981, 1986, 1991, 1996	기본형*
	2001	기본형*, 확장형
Child Care Arrangements Survey	1984	기본형
Child Care Survey	1993,1996	기본형*
	1999	기본형*, 확장형
	2002, 2005	확장형
Disability, Ageing and Carers	1993, 1998, 2003	기본형*
Domiciliary Services, Victoria	1986	기본형*
Education and Training Experience	1989, 1993, 1997	기본형
Education and Training	2001	기본형
Education and Training	2005	기본형*, 확장형
Education and Work	2001, 2003, 2005	기본형*
Families in Australia	1992	기본형
Family Characteristics Survey	2003	확장형
Forms of Employment	1998	기본형*
General Social Survey	2002, 2006	기본형*, 확장형
Giving to Charities Population Survey Monitor	1997	기본형
Household Expenditure Survey	1975-76, 1984, 1988-89, 1993-94	기본형*

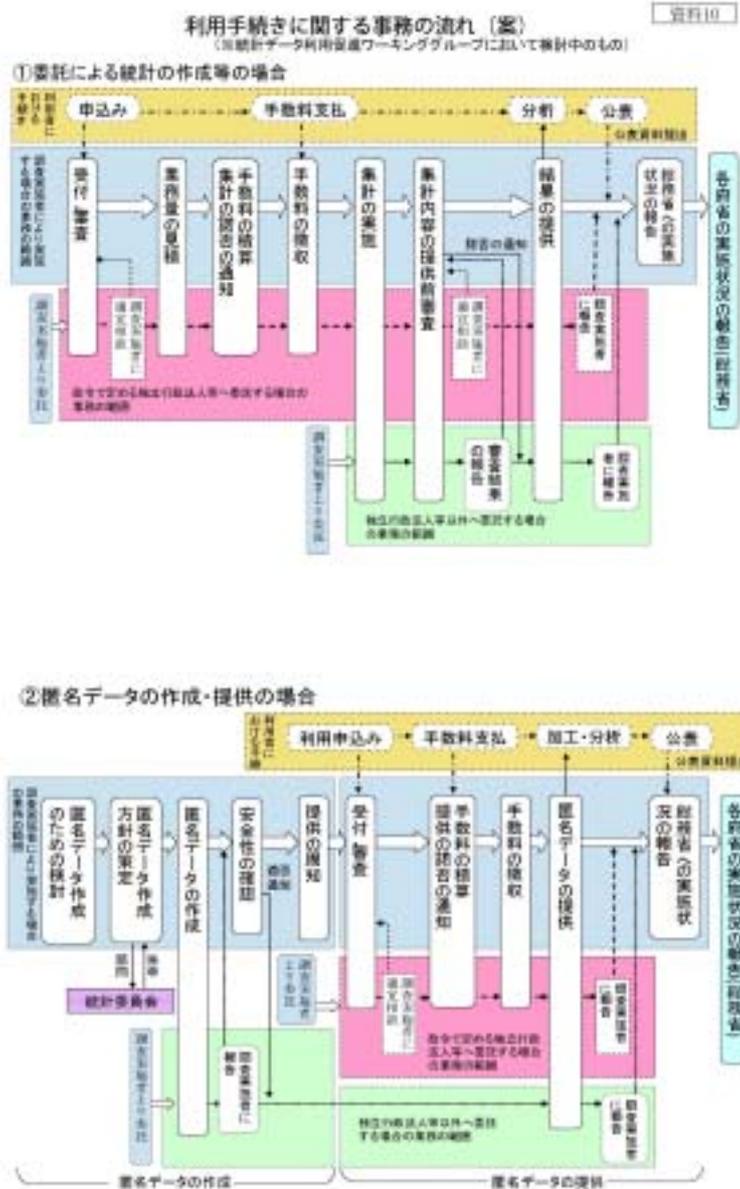
(계속)

조사명	기준연도	세분화 정도
Household Expenditure Survey and the Survey of Income and Housing (HES/SIH) Second Edition	2003-04	기본형*, 확장형
Income and Housing Survey	1981-82	기본형*
Income Distribution Survey	1986	기본형*
Income and Housing Costs Survey	1990	기본형*
Income and Housing Costs Survey	1994-98, 1999-2000	기본형*
	2000-01	기본형*, 확장형
Income and Housing Survey	2002-03, 2005-06	기본형*, 확장형
Income and Housing Survey and the Household Expenditure Survey (HES/SIH) Second Edition	2003-04	기본형*, 확장형
Innovation in Australian Business	2003	확장형
Labour Force Survey and Employee Earnings, Benefits and Trade Union Membership	2004	기본형*
	2006	기본형*, 확장형
Labour Mobility	1991, 1994	기본형*
Labour Force Survey and Labour Mobility Survey	2006	기본형*
Mental Health and Wellbeing of Adults, Australia, Second Revised	1997	기본형*
Multi Purpose Household Survey	2004-05, 2005-06	확장형
National Aboriginal and Torres Strait Islander Health Survey	2004-05	기본형*, 확장형
National Aboriginal and Torres Strait Islander Survey	1994	기본형*, 확장형

(계속)

조사명	기준연도	세분화 정도
National Aboriginal and Torres Strait Islander Social Survey	2002	확장형
National Crime and Safety Survey	2002, 2005	확장형
National Health Survey	1977-78, 1983, 1989-90, 1995	기본형*
	2001	기본형*, 확장형
National Health Survey, Indigenous	2001	확장형
National Nutrition Survey	1995	기본형*
Personal Safety Survey	2005	확장형
Pregnancy and Employment Transitions	2005	확장형
Rental Investors' Survey	1993, 1997	기본형*
Rental Tenants Survey	1994	기본형*
Superannuation and Employment Arrangements Survey	2000	확장형
Time Use Survey	1992	기본형*
	1997	기본형*, 확장형
Voluntary Work Survey	2000	기본형*
Women's Safety Survey	1996	기본형*

E. 일본의 자료 이용 흐름도(안)



## F. 각국의 마이크로데이터 제공비용

미 국	캐나다
<ul style="list-style-type: none"> <li>· Public Use File : 무료</li> <li>· RDC : 약 \$15,000/년 (기관 vs 기관)</li> <li>※ ICPSR<sup>1)</sup> (가맹자격에 따른 연회비)               <ul style="list-style-type: none"> <li>○ Category A : \$12,730</li> <li>○ Category B : \$7,930</li> <li>○ Category S : \$3,980</li> <li>○ Category C : \$2,000</li> </ul> </li> <li>- Category A : 사회과학 및 관련 분야의 대학원이 있는 대학 및 연구기관 (연구원 60명 이상)</li> <li>- Category B : 사회과학 및 관련분야의 대학원이 없는 대학 (학생수 2,500명 이상)</li> <li>- Category S : 사회과학 및 관련 분야의 대학원이 없는 대학 (학생수 2,500명 미만)</li> <li>- Category C : 개발도상국의 기관</li> </ul>	<ul style="list-style-type: none"> <li>· Public Use File : 파일마다 요금을 징수 (\$1,000 ~ \$3,000 정도) 일괄구입시 할인이 있기도 함</li> <li>· RDC               <ul style="list-style-type: none"> <li>- Direct Access : \$3,500/월</li> <li>- Indirect Access : \$5,500/월</li> </ul> </li> <li>※ DLI(Data Liberation Initiative)               <ul style="list-style-type: none"> <li>- 규모가 큰 기관(27개 기관) : \$12,000/년 (CARL, Canadian Association of Research Libraries 등)</li> <li>- 그 외의 기관 : \$3,000/년 (CASUL, Canadian Association of Small University Libraries 등)</li> </ul> </li> </ul>

주: 1) the Inter-University Consoritum for Political & Social Research.

(계속)

호주	영국	그 외 유럽
<ul style="list-style-type: none"> <li>· CURF : 데이터 세트당 약 \$1,000 ~ \$2,000</li> <li>· 컨설팅비용 : 프로젝트당 \$435</li> </ul>	<ul style="list-style-type: none"> <li>· CCSR<sup>1)</sup> (1991, 2001 SAR)</li> <li>- 학술목적 : 무료</li> <li>- CD-ROM 구입 : £30</li> <li>- 상업목적: 데이터 세트당 £1,000</li> <li>- 지방정부, 보건기관, 중앙정부, 자선단체 : £500</li> <li>· UKDA<sup>2)</sup> (사회, 인문 관련 서베이 데이터)</li> <li>- 학술목적 : 무료</li> <li>- 상업목적: 데이터 세트당 £500</li> <li>- CD-ROM: 기본 £7.5, handling fee £2.5, 우송비 £3(영국내), £4(유럽), £5(그 외)</li> <li>- 부가가치세 : 17.5%</li> <li>· CeLSIUS<sup>3)</sup></li> <li>- 국내 학술기관의 교직원, 학생에게 무료 제공</li> <li>· ONS-IDBR<sup>4)</sup></li> <li>- 약 £60/hour</li> </ul>	<p>&lt;이탈리아&gt;</p> <ul style="list-style-type: none"> <li>- 데이터 세트당 €90 (센서스자료 : €150)</li> <li>대학, 공공학회는 20% 할인</li> </ul> <p>&lt;독 일&gt;</p> <ul style="list-style-type: none"> <li>- €65 ~ €95</li> </ul> <p>&lt;네덜란드&gt;</p> <ul style="list-style-type: none"> <li>- 원칙: 무료</li> <li>- 실비보상</li> </ul>

주: 1) the Centre for Census and Survey Research.

2) the United Kingdom Data Archive.

3) Centre for Longitudinal Study Information and User Support.

4) Inter-Departmental Business Register.