제6장

센서스의 공간통계 생산과 제공에 있어서의 GIS의 역할

최은영

제1절 서 론

전수 조사인 센서스에서 조사 대상의 누락와 중복을 방지하기 위해서는 조사원에게 정확한 조사 범위를 할당하는 것이 중요한데, 이 때문에 정확한 조사구 경계가 표시된 조사구요도는 오래전부터 센서스 수행에 중요한 역할을 해왔다.[1] 과거에는 이러한 조사구 설정용 지도 제작에 도엽단위의 종이 지도를 복사하여 접합한 청사진 지도가 사용되었는데(통계청 2008b), 이는 가위와 테이프로 전국을 커버하는 거대한 조각 그림을 맞추는 것에 비유되는 지난한 과정이었다(Marx, 1986).

보다 많은 공간 정보가 생산되고, 정보 기술과 컴퓨터가 발달함에 따라 등장한 지리정보시스템(Geographic Information System: GIS)은 다양한 공간 단위의 통계 생산과 처리를 지원하고, 통계 데이터 수집 및 제공 등 센서스 전반에 걸쳐 다각도로 활용되고 있다.[2] 통계구역의 디지털 경계・각종 통계 주제도・공간 분석 기능의 제공과 인터넷 GIS를 통한 공간 통계의 제공 등은 이와 관련된 GIS의 주요 기능이다. 많은 비용과 시간・인력이 소요되는 대규모 프로젝트인 센서스의 효율적인 조사 수행과 고품질의 센서스 조사 결과 생산을 위해, 미국을 비롯한 유럽의 여러 국가들에서 GIS와 통계가 결합된 통계지리시스템은 필수적인 것으로 인식되고 있다(Peters and MacDonald, 2004). 기존의 통계 시스템과 GIS의 결합은 한쪽의 일방적인 필요에 의해서가 아니라, 상호간의 효용성이 증대된다는 점에서 자연스러운 현상이다. GIS는 각종 통계 업무를 효과적으로 지원

1) 지도는 공간 관계를 효과적으로 전달하기 때문에, 거처 등을 조사구에 할당하고, 조사구를 별도의 집계 단위로 재분류하는 센서스 수행을 위한 두 가지 기본 작업에 필수적인 요소가 된다.

2) GIS는 지리정보를 효율적으로 저장하고, 조작하고, 분석하기 위해 디자인된 정보 시스템으로 본 연구에서 사용된 개념들을 중심으로 GIS의 주요 개념을 <부록 1>로 정리하였다.

하며, 통계 데이터는 GIS 시스템에 풍부한 인구・사회학적 속성 자료들을 제공한다. 또한 최근 GIS와 더불어, 원격 탐사(remote sensing), GPS(Global Positioning System), 모바일 기기를 통한 위치인식서비스 등과 같은 새로운 기술이 발전함에 따라, 통계 조사・집계 과정에서 사용되는 전통적인 지도의 역할을 넘는 새로운 지리적 재현(representation)의 가능성이 커지고 있으며, 인터넷 기반의 지도 툴은 통계 정보의 도달 가능성을 확대시키고 있다(UN, 2008).

GIS의 발달과 함께 우리나라의 센서스에는 국립지리원의 종이지형도 원도를 스캔하여 제작된 스터 지도가 사용되었고, 이후 수치지도가 사용되었다.[3] 우리나라 통계청에서는 센서스 수행 전반에 걸쳐 GIS를 활용하고 있는데, 조사구 설정이라는 자료의 생산 단계뿐 아니라 집계구 설정, 통계지리정보서비스(SGIS) 등 자료의 제공 단계에도 GIS가 활용되고 있다.

우리나라에서는 GIS의 활발한 사용에 비해 논리적・이론적 연구는 거의 이루어지지 않았는데, 이는 미국과 영국을 중심으로 국가 통계 기관의 센서스 GIS에 관한 논문이 수백편 이상 출판되었고, 단행본(Peters and Macdonald, 2004; UN, 2008)이 출판되고 있는 해외의 상황에 비추어 볼때 이례적이다. 우리나라 통계청의 대내외적인 상황은 센서스의 공간통계 생산과 제공에 있어서의 GIS의 역할에 관한 이론적・논리적 분석을 요구하고 있다. 우선 가장 본질적으로는 보다 향상된 센서스 조사 결과를 얻기 위해 센서스 수행에 있어서의 GIS의 역할에 대한 연구가 이루어져야 한다. 이는 센서스 GIS의 발전이 보다 고품질의 센서스 결과를 얻기 위한 노력의 과정에서 이루어졌기 때문이다. 미국 센서스국 지리실(Geography Division)은 데이터의 수집・처리・집계를 향상시키는 혁신적인 방법을 고안하는 과정에서 타이거(TIGER)로 대표되는 수치화된 공간데이터베이스를 구축하였다. 이에 본 연구에서는 우선 센서스국 지리실의 역할이 GIS의 발달에 어떻게 공헌하였는지를 살펴본다.

통계청 내부적으로는 등록센서스로의 전환이 검토되고 있는데, 센서스에서의 GIS의 역할은 센서스 조사 방법론에 따라 변화한다. 이에 전통적인 현장 조사 방식의 센서스를 실시하고 있는 국가는 물론 등록센서스가 실시되고 있는 국가의 센서스 GIS에 관해서도 살펴볼 필요성이 제기되고 있다. 이와 관련되는 질문을 구체적으로 제기해보면, 첫째, 현장 조사 방식의 센서스 실시 국가에서는 GIS가 센서스 자료의 생산과 제공에 어떻게 활용되고 있는가? 둘째, 등록부에 기반하여 센서스 자료를 생산함으로써, 현장 조사가 이루어지지 않는 등록센서스 실시 국가에서는 GIS가 센서스의 자료 생산과 제공에 어떻게 활용되고 있는가? 현장 조사 방식의 센서스가 실시되고 있는 국가에 대해서는, 우선

3) 우리나라 통계청에서는 2006년 7월에 통계지리정보과가 창설되었으며, 2010년 1월 공간정보서비스팀으로 명칭이 변경되었다.

센서스 수행 단계별로 GIS의 역할을 살펴보았으며, 미국・영국・한국 센서스에서의 GIS 활용 사례를 구체적으로 살펴보았다. 등록센서스 실시 국가에서의 GIS의 역할은 국가별 지리통계시스템의 구축 배경과 활용 사례를 중심으로 살펴보았다.

다음으로 대외적으로는 국가지리정보의 통합 움직임이 존재한다. 우리나라에서는 지리정보를 국가의 중요한 사회간접자본으로 인식하여, 1995년 이후 국가 주도의 NGIS(National Geographic Information System) 사업을 추진하였다(<부록 3> 참조). 이에 여러 정부 부처와 지자체에서 NGIS 사업의 일환으로, 또는 개별 지자체 수준에서 지리정보를 생산・관리・활용 중에 있다. 하지만 이러한 지리정보시스템의 개별적인 운영은 중복 투자로 인한 예산 낭비 및 관련 인력의 비효율적 운용 등의 문제를 발생시켜, 현재 통계청 외부에서는 국가지리정보 통합이 이루어지고 있으며, 관련 논의가 지속적으로 이루어지고 있다.[4] 이에 따라 통계청에서 수행되고 있는 센서스 GIS의 특수성과 전문성에 대한 논리적・이론적 기반의 강화가 필요할 것이라 판단된다.

한편, 정책 입안자, 기업가, 도시계획가, 연구자, 학생 등 센서스 자료의 이용자들은 다양한 지리적 스케일에 따른 통계 데이터를 요구하고 있는데, 특히 최근에는 소지역 통계에 대한 수요가 증가하고 있다. GIS가 제공할 수 있는 데이터 통합 능력, 즉 다른 영역으로부터 다양한 정보들을 연계하는 능력은 소지역 통계 수요에 대해 국가 통계 기관이 즉각적으로 대처할 수 있도록 지원하는데, 해외의 여러 국가 통계 기관에서는 소지역 통계 수요에 대한 대응으로 격자망 통계를 생산・제공하고 있다. 우리나라 통계청에서도 소지역 통계 제공을 위해 공간 데이터베이스 구축과 서비스를 위한 개선 작업을 지속적으로 수행하고 있는데, 이러한 노력의 일환으로서 우리나라 통계청의 자료 구축 현황에 기반해서 격자망 통계의 제공 가능성을 살펴보고, 격자망 통계 생산 방법을 검토해 보았다.

제2절 미국 센서스 조사 방법의 변경과 GIS의 발달

1. 센서스가 GIS의 발달에 미친 영향

초기의 GIS는 1960년대 중반에 캐나다 전국의 토지이용을 전산화하기 위해 개발된, Canada Geographic Information System(CGIS)인데, 토지 면적의 측정에 이용되었다. 유사한 시기에, 미국 센서스국에서는 1970년 센서스 수행을 위해, 미국 도로의 디지털 정보

4) 구 건설교통부의 측지 분야의 GIS와 구 해양수산부의 해양 GIS, 행정자치부의 지적 관련 GIS가 국토해양부의 GIS로 통합되었는데, 국토해양부(2008)의 보고서는, 부처별로 분산된 측지・지적 업무가 통합에 따라 국가지리정보의 생산・관리・유통을 효율적으로 수행할 조직의 개편 및 통합 필요가 있음을 주장하고 있다.

를 생성한 DIME(Dual Independent Map Encoding)을 개발하였다. DIME 파일은 속성 정보로 저장된 주소만을 가지고 공간위치를 확인하고 지도화할 수 있도록 설계되었다. 즉 이 파일은 도로정보를 토대로 위치 정보를 자동으로 산출해주는 지오코딩(geocoding) 기능을 수행할 수 있도록 개발되었으며, 이후 보다 진보된 형태의 TIGER(Topologically Integrated Geographic Encoding Referencing) 파일 개발을 가능하게 하였다.[5)]

미국 센서스국에서는 수십 년 동안 GIS를 통계 조사와 서비스에 활용해 왔으며, 데이터베이스에서의 위상 구조 개념의 도입, 도로망 기준의 지오코딩 기법의 개발, 국가적 규모의 공간 데이터베이스 구축 등 GIS의 역사 발달에 중요한 역할을 담당한 것으로 평가받고 있다(Coppock and Rhind, 1991; Cooke, 1998). 센서스 조사 과정에서 구축된 DIME이나 타이거와 같은 공간 데이터는 GIS 데이터 구축에서 획기적인 획을 그었다. ESRI社나 Harvard Lab이 GIS 관련 소프트웨어 측면에서, 데스크톱 PC가 하드웨어 측면에서 중요한 역할을 했다면, 센서스국은 데이터베이스 측면에서 GIS의 발달에 크게 공헌하였다.

미국 센서스국이 GIS 발달에 기여할 수 있었던 이유를 알기 위해서는, 왜 센서스국이 이 분야를 다루고 있는지가 우선적으로 규명되어야 한다. 센서스국은 국회・행정부・기업가・학문 사회・일반 대중에게 인구와 경제에 관한 기본 정보를 제공해야 할 의무를 가지고 있는데, 이를 위해 센서스 데이터의 수집・처리・집계를 향상시킬 수 있는 혁신적인 방법을 고안할 필요가 있었다. 센서스의 성공적인 수행은 얼마나 자료를 잘 수집하느냐에 달려 있을 뿐 아니라 이러한 자료를 지리적 영역과 얼마나 잘 연결시키느냐에 달려있다. 이러한 노력의 결과는 수치화된 공간 데이터베이스의 구축으로 이어졌다. 센서스에 대한 지리적 지원 업무는 데이터 생산 및 제공의 효율성과 적시성 증가, 데이터의 질 향상, 비용 절감, 데이터 사용자를 위한 새로운 산출물 제공 등을 목표로 하고 있다(Sperling, 1995). 자료 수집과 처리를 위해서는 첫째, 거처를 블록 등의 공간 단위에 할당하고, 둘째 할당된 지역을 집계 단위로 분류해야 하는 두 가지 기본적인 작업을 수행해야 한다.[6)] 센서스국의 지리실(Geography Division)에서는 이를 지원하기 위해 거처에 정확한 지리적 위치를 할당하고, 할당된 위치를 집계구역에 맞게 집계할 수 있는 지도 및 지리적 산출물을 제공해야 했다. 지리실 업무의 범위와 복잡성은 센서스국의 센서스 방법론의 변화에 따라 달라졌다(Marx, 1986). 다음에서는 센서스국의 활동이 어떠한 산출물을 만들어 냈고, 그 산출물들이 어떻게 GIS 발달과 연관되었는지를 역사적으로 살펴본다.

---

5) 지오코딩 기능은 강력한 지도위치 입력 기능으로, 주소정보만 가지고 공간 위치 정보를 입력할 수 있기 때문에 다양한 GIS 기능 및 적용을 가능하게 한다(신정엽・이상일, 2008).

6) 미국 센서스 공간 단위의 표준 체계는 <부록 2>에 정리하였다.

2. 1990년 이전 센서스의 자료 수집 방법 변화와 지리적 지원

가. 전통적 방식의 1950년과 이전 센서스

1) 그림으로서의 종이 지도

1950년대 이전 센서스에서는 조사원(enumerator)들의 개별 가구 방문에 의해 데이터가 수집되었다. 조사의 중복과 누락을 피하기 위해 조사원들의 업무 할당을 지리적으로 조직하는 것이 필요했는데, 이를 위해 조사 지역의 범위와 특성을 기술해 주는 특수한 종이 지도가 개발되었다. 이러한 지도는 그래픽과 그림 형태로 지표를 기술했는데, 도로・철도・하천 및 다른 주요한 지표 형상들을 표현했다. 이외에도 센서스 단위 지역의 수치코드, 행정경계 등이 표현되어 있어 조사원은 이러한 정보를 바탕으로 주택의 위치를 지리적으로 할당했다.

[그림 6-1]은 당시에 사용되었던 지도의 전형적인 사례로 집계구(Enumeration District: ED)[7], 블록 도로 등의 다차원 관계가 표현되어 있다. 지도상에 나타나 있는 다차원 관계의 일부를 다음과 같이 지도가 아닌 기술을 통해 나타내 보면, 지도상에 얼마나 많은 공간 정보들이 표현되어 있는지를 알 수 있다.

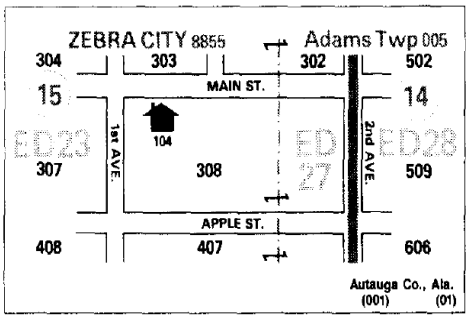
> 여기에서 2번가는 센서스 트랙(census tracks)의 경계이다. 센서스국은 이 센서스 트랙의 경계와 제브라시(Zebra city)의 경계를 집계구 할당에 사용했다. 15번 센서스 트랙은 2번가의 서쪽에 있으며 14번 센서스 트랙은 2번가의 동쪽에 있다. 28번 집계구(ED28)는 2번가의 오른쪽에 있고, 27번 집계구(ED27)는 2번가와 시 경계 사이에 있고, 23번 집계구(ED23)는 시 경계의 서쪽에 있다.....

조사 지역이 할당되면 조사원은 이 지도를 조사구 내의 블록들에 대한 안내서로 사용한다. 주소 명부(address register)에 블록을 따라 있는 모든 주택의 목록을 기입하고, 개별 주택이 위치해 있는 블록의 번호를 적는다. 23번 집계구에 할당된 조사원은 메인가(Main St.)를 따라 걸으면서 그곳에 있는 모든 개별 주택의 명부를 작성하는데, 메인가 104번지는 308번 블록에 위치해 있다고 명부에 기입한다. 이와 같이 조사원은 블록의 번호를 적는 간단한 방법을 통해 수작업에 의한 일종의 지오코딩 기능을 수행함으로써, 주택을 특정 지역에 할당하는 첫 번째 지리적 지원 기능을 완수한다.

지표에 대한 '그림'으로서의 지도는 도로, 광범위한 지리적 정보, 개별 단위들의 관계를 보여준다. 그러나 조사원이 이 지도를 사용할 때는 도로와 집계구역 사이의 다차원적

7) 대략 인구 800~1000명 사이의 농촌 지역이나 인구 희박 지역을 가르키는 미국 센서스국의 공간 단위이다.

인 모든 관계가 아닌 위에서 기록된 가장 기본적인 지리적 정보만을 사용한다. 따라서 조사 과정에서 중복과 오류가 발생할 가능성이 크다.



Serial Number	Address (House Number and Street Name)	Block Number	ED Number	Occupant Name(s)
0001	104 Main Street	308	23	James
0002				
0003				
0004				

출처: Marx, 1986.

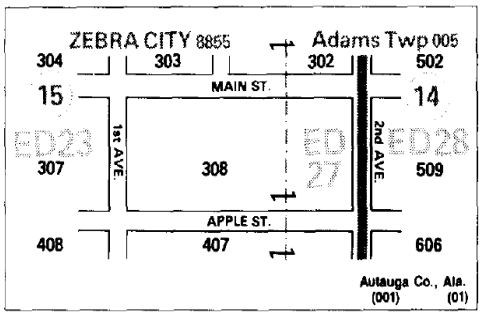
[그림 6-1] 센서스 지도와 주소 명부

2) 집계의 기반으로서의 지리 참조 파일

조사원이 지도를 사용하는 방식을 보완하고, 두 번째 지리적 지원 기능을 수행하기 위해 각각의 공간 단위들은 센서스의 모든 집계 단위에 할당된다. 이를 위해 지리실은 또 다른 특별한 수단을 개발했는데, 이것이 지리 참조 파일(geographic reference file)이다. 이 파일은 미국 전역의 모든 공간 단위들 간의 관계를 보여준다.

[그림 6-2]는 전형적인 조사구요도의 지리 참조 파일의 일부이다. 각각의 집계구에는 주, 카운티, MDC/CCD, 장소, 센서스 트랙, 블록 코드가 할당되어 있다.[8] 23번 집계구는 제브라시 안에 있고, 15번 센서스 트랙 안에 있으며, 27번 집계구는 제브라시 밖에 있지만 15번 센서스 트랙 안에 있다. 28번 집계구는 제브라시 밖에 있고 14번 센서스 트랙 안에 있다. 이러한 지리적 관계 정보를 이용해서 센서스국은 조사원에 의해 명부에 기입된 모든 주택을 집계 구역에 할당한다. 예를 들어 조사원에 의해 308 블록이라고 할당된 메인가 104번지는, 01번 주(알라바마), 001번 카운티(Autauga), 8855번 장소(제브라시), 15번 센서스 트랙으로 할당된다. 이상에서 살펴본 1950년대 이전의 센서스 방법론과 지도의 역할을 도식화해서 살펴본 결과는 다음과 같다([그림 6-3]).

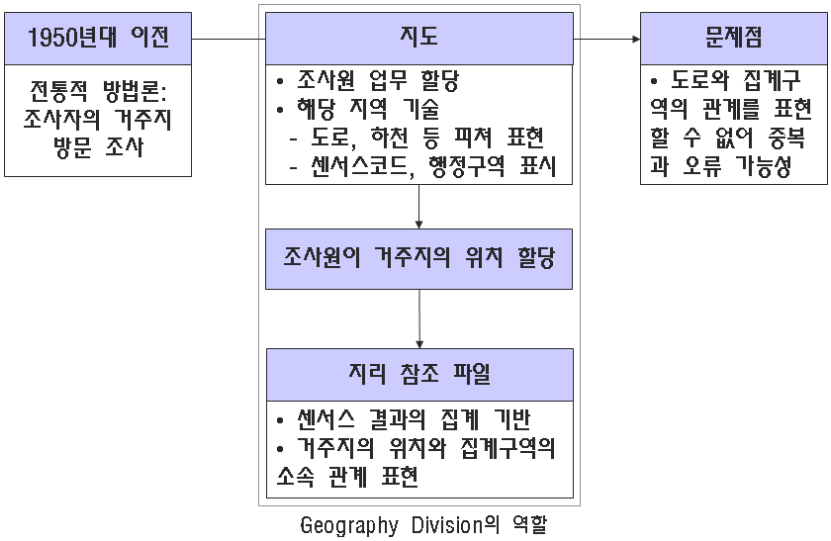
8) MCD(Minor Civil Divisions)는 법적 단위이며, CDC(Census Counry Divisions)는 통계 단위인데, 카운티보다 작은 동일한 수준의 공간 단위이다.



	State	County	MCD/ CCD	Place	Census Tract	Block Numbers
ED23	01	001	005	8885	15	302P 303 304 307 308P 407P 408
ED27	01	001	005	–	15	302P 308P 407P
ED28	01	001	005	–	14	502 509 606

출처: Marx, 1986.

[그림 6-2] 센서스 지도와 지리 참조 파일

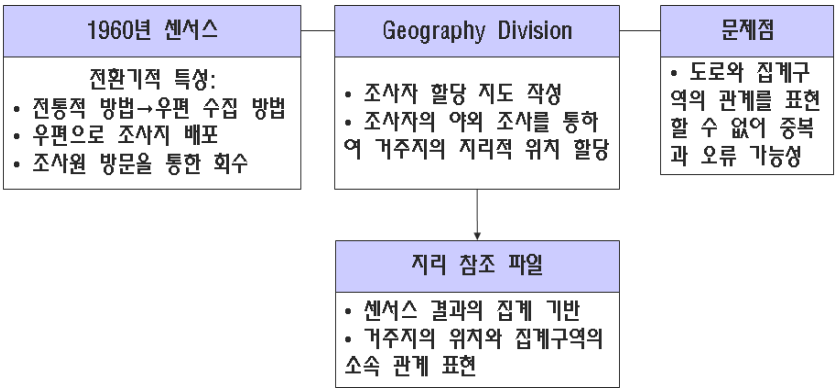


[그림 6-3] 1950년대 이전의 센서스 방법론과 지도의 역할

나. 전환기의 1960년 센서스

1960년 센서스의 데이터 수집 방법은 전통적인 방식에서 우편 조사로 부분적으로 변경되었다. 조사 방법의 변경은 조사표 배부에 국한되었는데, 조사표 배포는 우편을 이용하되, 조사표 회수는 종전과 같이 조사원이 가구를 방문하는 방식으로 이루어졌다. 이러한 방법론의 변경에도 불구하고 1960년 센서스에서 지리실의 역할은 1950년 센서스와 동일하였다. 지리실은 조사원의 할당을 위한 지도를 작성하였고, 조사 기간 동안 개별 주택의 지리적 위치를 할당하는 작업은 전적으로 조사원의 현장 조사에 의해 이루어졌

다. 지리 참조 파일 또한 여전히 각 위치를 적절한 집계 단위로 분류하기 위하여 사용되었다. 1960대의 센서스 방법론과 지도의 역할을 도식화해서 살펴본 결과는 다음과 같다([그림 6-4]).



[그림 6-4] 1960년 센서스 방법론과 지리실의 역할

다. 지리적 업무의 초기 자동화가 이루어진 1970년과 1980년 센서스

1960년대 중반에 센서스국은 향후 센서스의 데이터 수집 방법을 우편으로 조사표를 배부하고 우편으로 회수하는 우편조사(mail-out/mail-back)로 결정하였다. 이러한 데이터 수집 방법의 변화와 지리적 업무의 자동화를 도입한 배경으로는 전통적인 개별 가구 방문에 의한 조사의 어려움과 비용의 증가, 번지와 도로 번호에 기초하여 우편배달이 가능한 도시지역 거주 인구의 증가, 주소와 집계를 위한 지리적 단위를 연계하는 컴퓨터 기술의 발달 등을 들 수 있다(Sperling, 1995). 이러한 변화에 맞추어 센서스의 수행 과정에서 지리실의 지리적 지원 활동도 변화해야 했다. 조사원은 더 이상 영역 할당 지도를 받아 가구를 방문할 필요가 없어졌으며, 현장 조사에 기반해서 개별 주택에 블록 코드를 부여하는 작업을 할 필요가 없어졌다. 이제 조사원이 해야 하는 일은 센서스 조사표를 제출하지 않은 가구를 여러 번 방문하는 것이다(Marx, 1986). 사람은 더 이상 지오코딩 정보의 유일한 자료원이 아니게 되었다. 이러한 방식의 성공은 지도를 지참한 조사원을 현장에 보내지 않고도 각 센서스 조사표를 적절한 센서스 블록에 '지오코딩' 할 수 있는지에 달려 있었다(Cooke, 1998).

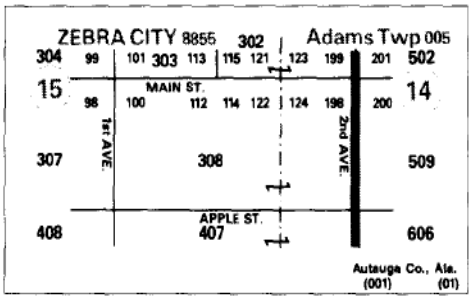
1) 자동화된 공간 데이터데이스를 향한 첫 걸음으로서의 ACG

과거에 조사원이 하던 작업을 지도가 대신하기 위해 지리실은 주택과 도로, 센서스

국 지도상의 또 다른 객체들의 관계에 관한 정보를 코드화하는 방법론을 고안했다. 1970년 센서스에서 이 새로운 지리적 방법은 Address Coding Guides(ACG)라고 불린다. ACG는 우편에 의해서 수집된 144개 대도시권의 우편 주소를 지오코딩하는데 필요한 지리기반 파일의 프로토타입(prototype)으로, 도로이름, ZIP 코드, 주소 범위, 주소 범위에 해당하는 센서스 트랙 및 블록 번호에 대한 정보를 담고 있다.

앞의 전통적인 지도로 살펴본 사례 지역의 ACG는 다음과 같다([그림 6-5]). ACG를 작성하기 위해서는, 센서스국의 지도와 국지적인 지리 참조 소스를 사용하여 도로가 교차될 때 생기는 각 세그먼트마다 도로 한쪽 편의 정보를 한 레코드에 입력해야 한다. 완성된 ACG의 표에는 도로명, 메인가와 교차하는 개별 블록에 포함되는 주소의 범위, 센서스 트랙, 블록과 같은 집계 지역의 코드와 같은 지도상에 표현되어 있는 기본적인 지리 정보들이 기재된다.

그러나 지오코딩의 기반으로서 ACG는 여러 가지 문제점을 갖고 있었다. 우선, ACG 파일은 센서스 지도상의 도로에 대한 정보만 포함할 뿐 철도, 하천 혹은 지도에 보이는 경계들에 대한 어떠한 정보도 포함하지 않았다. 또한 ACG의 파일 구조는 위상구조(topology)에 기반을 두고 있지 않기 때문에 지리실은 인코딩 과정에서 발생하는 많은 오류를 확인하거나 수정할 수가 없었다.[9] 마지막으로 ACG는 공간 및 위치 정보를 포함하지 않아 이를 기반으로 컴퓨터를 이용하여 지도를 그릴 수가 없었다(Marx, 1986).



Street Name	Place	Census Tract	Block Number	Address Range
MAIN ST.	8855	15	303	101-113
MAIN ST.	8855	15	302	115-121
MAIN ST.	----	15	302	123-199
MAIN ST.	8855	15	308	100-122
MAIN ST.	----	15	308	124-198

출처: Marx, 1986.

[그림 6-5] 센서스 지도와 ACG의 데이터 구조

2) DIME

지리적 데이터 처리에 대한 센서스국의 실질적인 참여는 1967년 New Haven Census

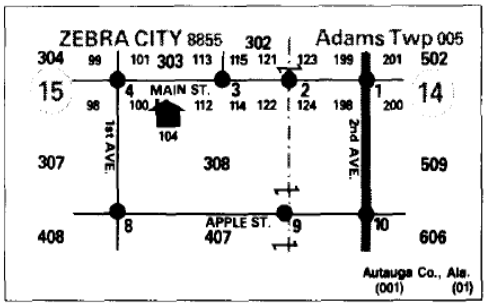
9) 인코딩은 컴퓨터에서 저장하거나 사용할 목적으로 문자나 숫자를 부호화하는 방법을 의미한다.

Use Study와 함께 시작된다. 이 연구는 센서스국의 지리실이 만들어낸 ACG에 기반해서 시작되었다. 이 연구에서는 어떻게 0, 1, 2 셀이 행렬을 통하여 연결될 수 있는지를 기술하는 공간 위상 구조에 대한 개념을 논의하기 시작하였다. 1967년 이러한 위상 구조를 바탕으로 DIME(Dual Independent Matrix(후에 Map) Encoding)이 탄생하게 된다. 여기에서 'Dual'은 두 개의 행렬, 즉 0/1 셀과 1/2 셀 행렬을 나타내고, 'Independent'는 센서스국의 지리실과 독립적으로 개발되었다는 의미이다. 실제로 1970년 센서스에서는 ACG만 사용되었다. 각 라인의 세그먼트는 두 개의 노드만을 가질 수 있고, 두 센서스 블록 사이에만 존재할 수 있다. 따라서 각 라인 세그먼트는 도로 이름, 시작점(From node), 끝점(To node), 좌측 블록, 우측 블록, 좌측 주소 범위, 우측 주소 범위, 좌측 ZIP 코드, 우측 ZIP 코드 등의 정보를 갖는다.

지리실은 1980년 센서스를 위하여 1970년 센서스에서 사용된 ACG를 일련의 Geographic Base File/Dual Independent Map Encoding(GBF/DIME) 파일로 변환하였다. 지리실은 기존 대도시권에 대한 업데이트와 모든 대도시 지역에 관해 GBF/DIME를 생성하는 프로그램을 지원하였다. GBF/DIME 파일은 ACG와 같이 기본 지리 정보를 포함하고 있을 뿐 아니라 대도시권 지도(Metropolitan Map Series)에 포함되어 있는 철도, 하천, 센서스 집계 경계 등 다른 피쳐(feature)들에 대한 정보도 포함하고 있었다. 더욱이 GBF/DIME 파일의 준 위상 구조(semi-topology)는 모든 교차점에 노드 코드라고 불리는 고유 식별자를 가지고 있는 작은 점을 지도에 추가시켰다([그림 6-6]). GBF/DIME 파일의 레코드는 지도상에 표현된 지리정보, 각 도로의 세그먼트 별 주소 범위, 이와 인접한 블록의 번호 등을 저장한다.

이 파일을 완성하는 과정에서 4번 노드에서 3번 노드, 3번 노드에서 2번 노드, 2번 노드에서 1번 노드로의 순서로 도로의 각 세그먼트에 관한 정보가 채워진다. 이 파일은 지도에 나타나 있는 지리적 정보를 보여주는데, 메인가의 각 교차점에 해당하는 주소의 범위, 인접 블록의 정보 등을 보여준다.

ACG와 GBF/DIME 파일은 모두 주소 참조 파일로, 주소 범위와 블록과의 관계 정보를 바탕으로 컴퓨터 매칭 알고리즘을 이용하여 어떤 주소가 어떤 블록에 속하는지를 파악하는 것을 가능하게 해준다. GBF/DIME 파일은 주요 핵심 대도시권과 비대도시권의 도시화된 지역을 커버한다. GBF/DIME 파일을 통하여 미국 전체 가구의 40~50%를 지오코딩하는 것이 가능하며, 이때 지오코딩의 정확도는 센서스 블록 수준에서 96% 정도이다 (Sobel, 1990).



Street Name	From	To	Place	Census Tract	Left Block	Right Block
MAIN ST.	4	3	8855	15	303 101-113	308 100-112
MAIN ST.	3	2	8855	15	302 115-121	308 114-122
MAIN ST.	2	1	——	15	302 123-199	308 124-198

출처: Marx, 1986.

[그림 6-6] 센서스 지도와 GBF/DIME 파일

라. 초기 시스템의 문제점

지금까지 살펴본 센서스에 활용된 초기 지리 시스템에는 문제점이 존재한다. 지도, ACG, GBF/DIME 파일 등의 지리 참조 파일을 포함한 과거 센서스로부터 만들어진 모든 지리적 산출물들은 지표를 기술하는데 있어 몇 가지 점에서 공통점을 가지기는 하지만, 지표를 기술하는 서로 다른 방법을 취하였다. 센서스를 준비하기 위하여 수많은 사람들이 이러한 자료를 다루게 되면서 동일한 산출물에 상이한 오류들을 만들어 냈고, 이는 최종 산출물에서의 불일치를 야기하였다.

앞에서 살펴본 [그림 6-2]는 지리 참조 파일을 필사하는 과정에서 나타나는 일반적인 오류 유형을 보여준다. 여기에서는 제브라시의 코드가 8855대신 8885라고 지리 참조 파일에 잘못 기입되어 있다. 지도와 참조 파일의 이러한 불일치는 이 도시와 관련된 일련의 지리적 산출물에 많은 문제를 야기하며, 이는 센서스국 직원들과 데이터 사용자의 불만을 야기하게 된다. 지도와 관련 자료들을 생성하기 위한 복잡하고 기능적으로 분리된 과정들은 여러 가지 문제를 야기하였다(Marx, 1986).[10] GBF/DIME 파일의 일반적인 약점 중 하나는 파일이 커버하는 공간의 범위가 대도시 지역에 제한된다는 점이다. 또한 상대적으로 소수의 GBF/DIME 파일만이 1980년대 동안 갱신되고 유지되었다. 갱신된 파일 중에서도 일부만이 센서스국의 표준을 만족시켰다(Sperling, 1995).

10) 1980년 센서스에서는 미국 전역을 커버하는 32,000개의 지도를 만들기 위해 인디아나 주 제퍼슨빌 센서스국의 자료 처리 사무실에 약 900명, 캘리포니아에 계약직으로 300명의 사람들이 고용되었다. 이들은 지도를 그리고, 가능한 한 모든 도로에 명칭을 부여하고, 경계를 그리고 250만 개의 블록 코드를 할당하는 작업을 수작업으로 수행하였다. 센서스 실시 이전 2년 동안 이 작업이 수행되었으며, 상당한 수의 인원 감축 이후 관련 지도의 수정 작업이 센서스 이후 2년 동안 이루어졌다.
32,000개의 종이 지도를 만든 센서스국의 여러 직원들은 복사본을 만들어서 가위와 테이프를 이용해서 개별 조사원들이 사용할 약 30만 장의 조사구 할당 지도를 만들었다. 이 과정은 겹치지도 않고 두 조각 사이에 비는 것도 없는 거대한 조각 그림을 맞추는 과정과 유사하였다.

3. 1990년 센서스 이후 지리적 지원 기능의 자동화와 타이거

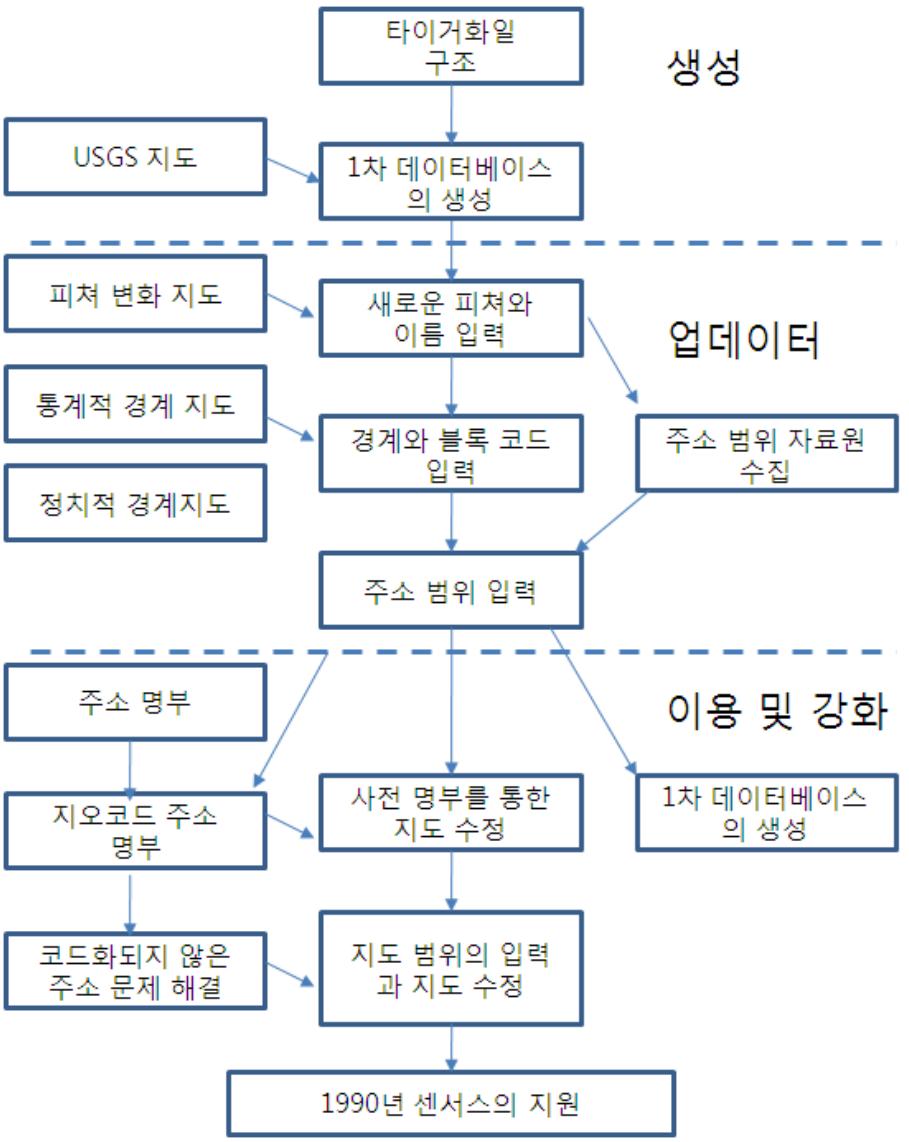
센서스국은 1990년 센서스를 위하여 이전 센서스의 성과물과 미국 지질조사국(USGS)의 1:100,000 축척 지형도를 바탕으로 타이거(Topologically Integrated Geographic Encoding Referencing: TIGER)라는 고유하고 새로운 파일 구조를 개발하였다. Marx(1986)는 이 시스템이 센서스와 각종 조사에 필요한 지도화와 관련된 지리적 활동들을 자동화할 것이라고 예상했다. TIGER에서 'topologically'는 지리적 영역을 정의하기 위하여 지도상의 점과 선이 어떻게 관련되는지에 대하여 과학적으로 설명하는 위상 구조를 의미한다. 타이거 시스템의 설계 과정에서는 지리적 구조에 대한 수학적 설명을 위하여 위상학, 그래프이론 및 관련 수학 이론이 수용되었다. 'Integrated'는 타이거 파일이 지표와 경계 피쳐들에 대한 자동화된 기술(description)을 통합하는 데이터베이스라는 것을 의미한다. 타이거 파일의 위상 구조는 도로, 하천, 철도 사이의 위치 및 관련성, 이들과 다양한 센서스 집계 영역 사이의 관련성을 정의한다. 즉, 필수적인 지리 정보들이 독립적으로 분리되어 존재하는 것이 아니라 단일한 전산 데이터베이스의 일부로 존재한다는 것을 뜻한다. 'Geographic'은 타이거 파일이 센서스국이 센서스와 조사를 수행하는데 사용하는 지리적 구조의 자동화된 기술을 저장하기 위한 저장소라는 것을 의미한다. 타이거 파일은 센서스 조사에 필요한 모든 지표상의 피쳐들에 대한 정보를 포함한다. 타이거 시스템의 주된 목적 중 하나는 이러한 피쳐들을 빠짐없이 그리고 중복 없이 구축하는 것이다. 'Encoding'은 지리 정보를 컴퓨터가 인식할 수 있는 형태로 재현하는 과정을 의미한다. 타이거 파일의 구축은 자동화된 지도 스캐닝, 디지타이징, 표준 데이터의 입력, 정교한 컴퓨터 파일 매칭 등 다양한 인코딩 기법을 통하여 이루어졌다. 마지막으로 'Referencing'은 미국의 지리정보에 대한 기준을 의미하는 것으로, 타이거 시스템은 센서스 관련 지리 정보에 대한 접근 및 검색 결과의 일관성을 보장함으로써, 이전 자료의 처리에서 센서스국과 데이터 사용자가 경험하였던 많은 혼란으로부터 벗어날 수 있게 한다(Marx, 1986). 요약하면 타이거 시스템의 목표는 센서스 지원을 위한 정확하고 일관성 있는 지도 기반을 구축하는 것이다. 한편 이러한 타이거 시스템은 단순한 데이터 저장소 이상의 의미를 갖는데, 타이거는 데이터, 데이터베이스관리시스템, 관리 및 조작을 위한 응용 프로그램, 문서 및 절차를 포함하는 '시스템'을 의미한다.

가. 타이거 시스템의 구축

1990년 센서스의 타이거 시스템은 과거의 문제들을 해결하기 위한 다양한 새로운 접근을 통하여 구축되었다. 미국 전역을 포괄하는 타이거 시스템의 개발 계획은 몇 가지 단계를 거쳐 이루어졌다. [그림 6-7]은 이러한 타이거 시스템의 구축 과정을 도식적으로 표현한 것이다. 구축 과정은 생성, 업데이트, 이용 및 강화 국면으로 나누어 살펴볼 수 있다.

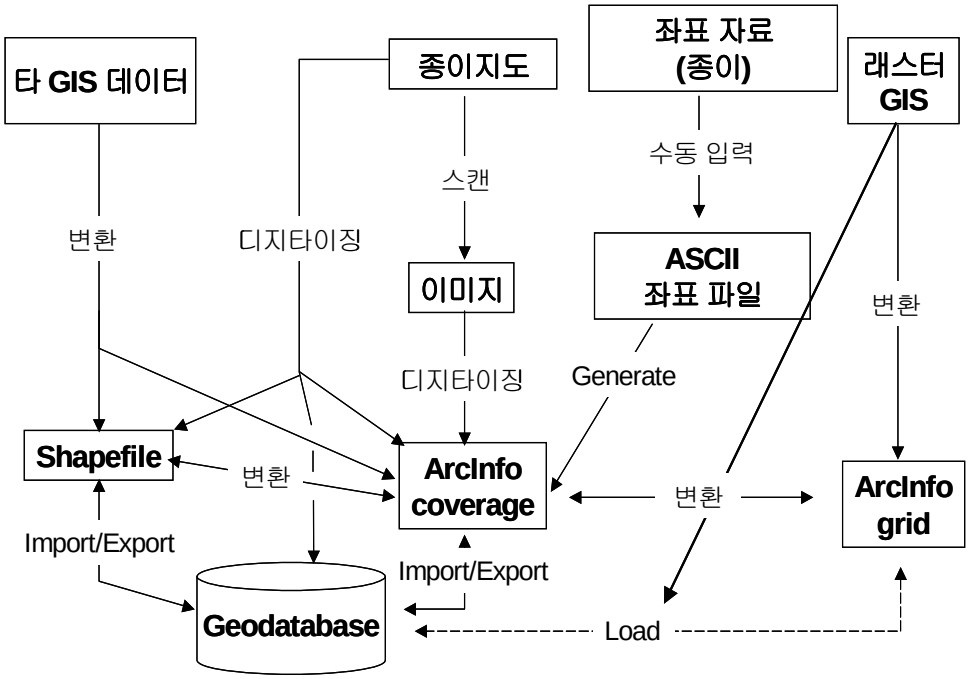
1) 1차 데이터베이스의 생성

GIS의 이용에 있어 공간 데이터의 획득은 전체 시스템 구축 비용의 60~85%를 차지하는 중요한 작업이다(이상일 외, 2009). GIS 데이터는 일반적으로 1차 자료와 2차 자료를 통해서 획득된다. 1차 자료원은 공간 데이터의 직접적 측정과 입력을 통해서 이루어지는데, 래스터 데이터의 경우, 인공위성 이미지나 항공사진 등을 통해, 벡터 데이터는 현장 측지나 GPS(Global Positioning System) 장비를 통해서 얻어진다. 2차 자료를 통한 데이터의 획득은 기존에 구축되어 있는 데이터들, 예를 들어, 종이 지도, 항공사진, 그 외 문서 자료들을 이용하여 새로운 데이터베이스를 생성하는 것을 말한다. 주로, 대형 스캐너를 통해 래스터 데이터가 생성되고, 테이블 및 스크린 디지타이징(digitizing), 입체 판독 등을 통해 벡터 데이터가 생산된다. 테이블 디지타이징의 기본 프로세스는 종이지도를 컴퓨터와 연결된 디지타이징 테이블 위에 부착한 후, 디지타이징 마우스를 지도상의 객체들을 따라 정확하게 클릭하여 좌표를 입력하는 것이다. 이러한 과정을 통해 입력된 좌표는 소프트웨어를 통해 컴퓨터에 자동으로 입력되며, 디지털화된 벡터 데이터를 생성하게 된다. 이러한 전 과정을 벡터라이징(vectorizing)이라 하며, 이는 초기 벡터 공간 데이터베이스 구축을 위한 가장 전형적인 방법이었다. 이후 고해상도의 대형 스캐너가 개발됨에 따라, 스캔된 디지털 이미지 파일을 가지고, 다소 공간을 차지하고 번거로운 작업이 필요한 테이블 디지타이징 대신에 컴퓨터의 스크린을 통해 같은 작업을 할 수 있게 되었다. 기존에 이미 디지털화된 공간 데이터베이스가 있는 경우에는 간단한 변환 툴을 이용하거나, 재가공하여 새로운 공간 데이터를 생성할 수도 있다. [그림 6-8]은 GIS 데이터 획득의 일반적 과정을 도식적으로 보여 주고 있다.



출처: Marx, 1986.

[그림 6-7] 타이거 파일 구축 과정



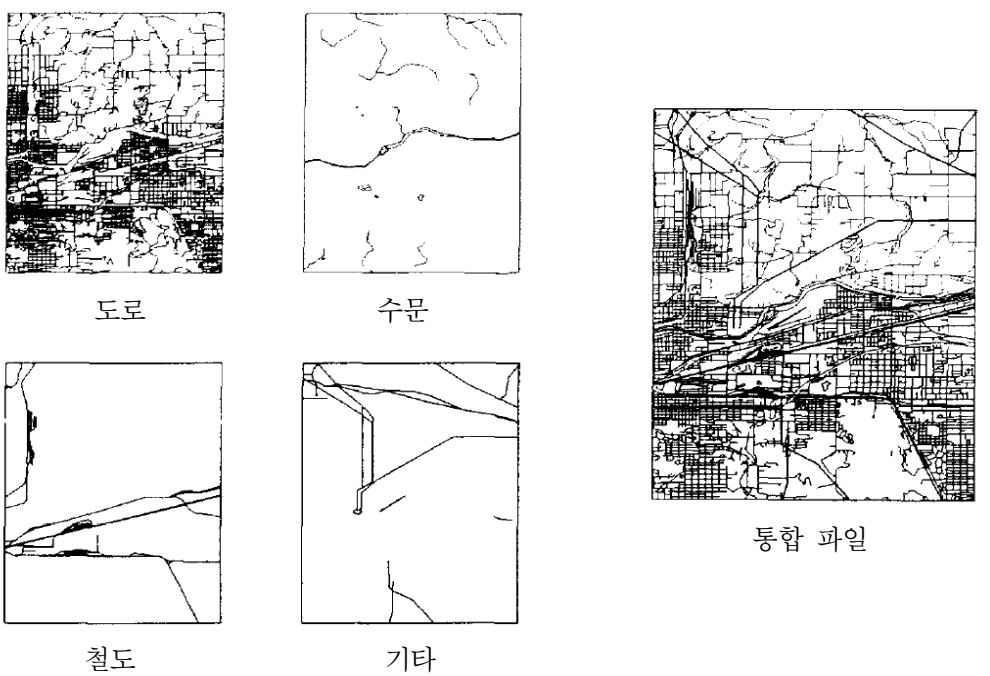
[그림 6-8] GIS 데이터의 획득 과정

타이거 파일의 개발에서 핵심은 정확하고, 일관성 있는 지도 기반을 구축하는 것이었다. 다른 연방기관, 주, 민간 부문에서 이루어지는 지리적 자동화 작업과의 중복을 피하기 위해 센서스국은 다수의 협정과 위원회에 참여하였다. 그 결과 센서스국은 지질조사국의 1:100,000 지형도를 컴퓨터가 읽을 수 있는 지도 파일로 변환하기 위하여, 지질조사국과 협력 프로젝트를 진행하였다. 지질조사국 외에도 타이거 구축 과정에 참여한 기관을 정리한 결과는 다음과 같다(<표 6-1>). 지질조사국은 지도에 나타나 있는 강, 도로, 철도, 전선, 파이프라인과 같은 피쳐들을 스캔닝한 후 벡터라이징을 통해 컴퓨터가 읽을 수 있는 형태로 변환하였으며, 이를 수직적으로 통합하였다([그림 6-9]). 도로를 제외한 '호수', '강' 등과 같은 피쳐들에 분류 코드를 할당하였다. 센서스국의 지리실에서는 도로 데이터에 피쳐 분류 코드를 할당하여 1차 데이터베이스를 생성하였다.

〈표 6-1〉 타이거 구축 과정의 참여 기관

프로그램 단계	USGS 100k 파일	GBF/DIME 파일	다른 지역의 센서스 파일
원파일의 디지타이징	지질조사국	센서스국	센서스국
피쳐 속성 명명	지질조사국・센서스국	-	센서스국
기존 피쳐의 이름과 우편번호코드의 확인	-	센서스국	-
수직적 통합	센서스국	-	센서스국
업데이트된 피쳐의 디지타이징	센서스국	민간 계약업자	센서스국
피쳐 이름 등록	센서스국	민간 계약업자	센서스국
새로운 도로에 대한 범위와 우편번호코드 입력	-	센서스국	-
지리적 코드 수정・입력	센서스국	센서스국	센서스국
타이거 파일로의 전환	센서스국	센서스국	센서스국

출처: Sobel, 1990.



출처: Carbauch and Marx, 1990.

[그림 6-9] 지질조사국 피쳐 파일의 수직적 통합

2) 데이터베이스의 업데이트

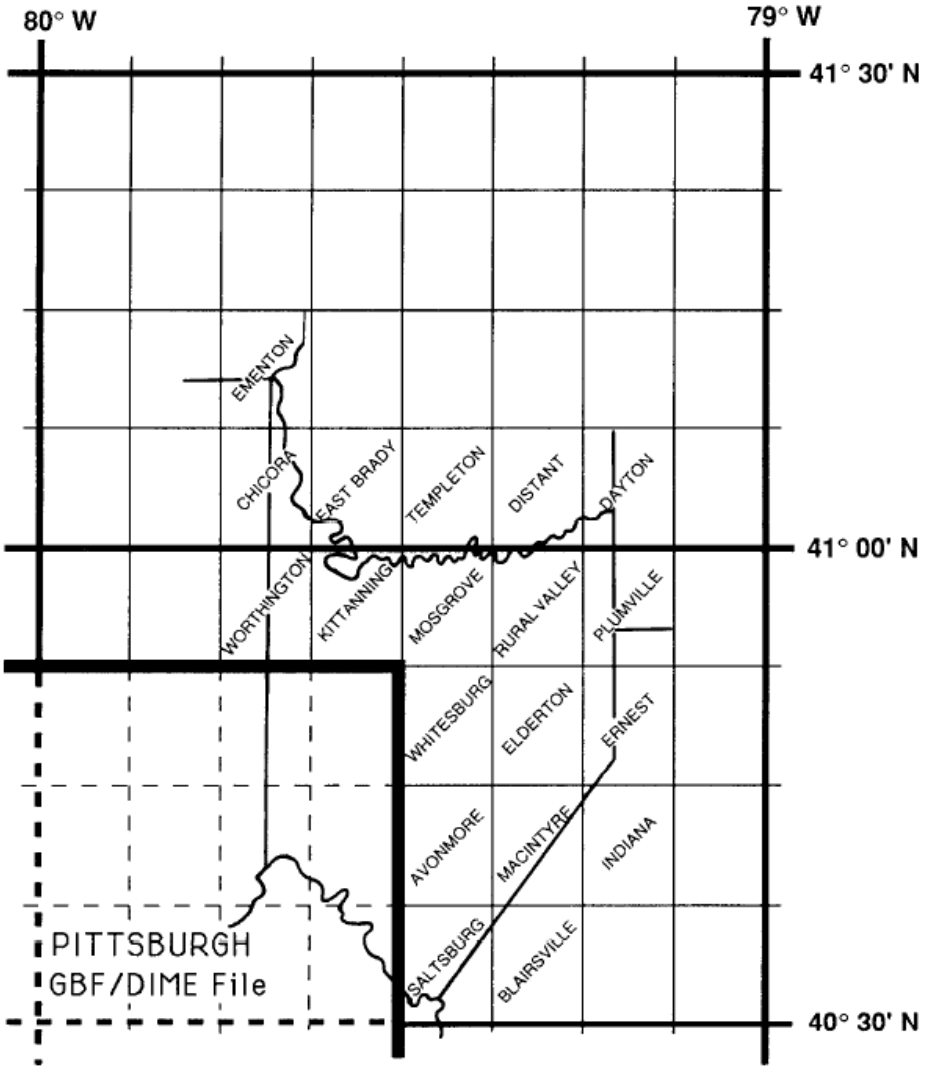
센서스국이 지질조사국으로부터 받은 데이터는 도로명이나 하천 이름이 없는 라인으로만 구성된 것으로, 그 자체로는 현장 조사와 사용자에게 유용한 형태의 지도는 아니었다. 따라서 이의 업데이트를 위한 일련의 작업이 수행되었다.

유용한 기본도를 만들기 위한 추가적인 정보를 모으기 위해 센서스국은 12개 지방 사무소마다 핵심 지리 지원 요원을 임명하였다. 이 요원들은 지도상의 피쳐들을 명명할 수 있는 자료원을 알아내고・평가하고・획득하기 위해, 그리고 새롭게 개발된 지역이 어디인지를 알기 위해, 주・카운티 등의 공무원들과 함께 공동 작업을 수행하였다. 이들은 미국 전역의 지방정부로부터 입수 가능한 최신의 지도를 수집하였으며, 새로 수집된 자료의 정보와 기존 시스템에 포함되어 있는 정보를 비교하였다(Carbauch and Marx, 1990). 차이가 있는 경우 누락된 도로나 도로명을 기입하고, 기존의 정보가 잘못된 경우에는 이를 수정하였다. 지리 요원들은 현장 디지타이징 담당(Field Digitizing Sites)으로 지정되어 있는 4개의 지방청(Regional Office)에 보내어지는 '피쳐 변화 지도(Feature Change Map)'에 수정이 필요한 정보를 표시하기 위해 수집된 자료들을 활용하였다. 이 작업에 4년 이상의 기간이 소요되었다(Carbauch and Marx, 1990). 그 결과 라인과 함께 이름을 가진 지도가 만들어졌다.

한편 12개 지방 사무소의 지리 요원들은 센서스국이 1990년 센서스의 자료 수집과 집계에 사용하게 될 통계적 경계를 정의하기 위해, 일련의 지방 공청회(local public meeting)에서 모아진 정보에 근거해서 주, 카운티, 지방 공무원들과 함께 공동 작업을 수행하였다. 지방청에서는 그 결과를 센서스국의 자료 처리 담당과에 전송하였다. 이와 함께 센서스국은 미국 전역의 카운티, 타운십, 도시의 경계 변화를 확인하기 위하여 연간 조사를 수행하였다. 제퍼슨의 본청에 있는 지리 요원들은 1980년 경계에 기반해서 지방청에서 올라온 통계 지역의 경계 변경에 관한 정보들을 처리하고, 업데이트된 디지털 형태의 지도 데이터베이스에 입력했다.

풍부한 정보를 담고 있지만 중간 정도의 좌표 정확도를 갖는 수치화된 GBF/DIME 파일은 1990년 타이거 데이터베이스에서 345개의 도시 지역에 대한 지도학적 기초를 형성하였다. GBF/DIME 파일의 개발에서 절대적 좌표 정확성은 일차적 관심의 대상이 아니었다. 왜냐하면 센서스국은 지도 작성 목적보다는 지오코딩의 목적을 위하여 GBF/DIME을 사용했기 때문이다. 초기에는 GBF/DIME 파일에 존재하는 데이터에 대한 손실 없이 지질조사국의 1:100,000 축척의 지형도와 결합하고자 하였으나, 시간과 논리적 이유로 이러한 시도가 비현실적이라는 것이 인식되었다. 대안으로 접근한 방법이 지질조사국의 1:100,000 지도의 도시 지역을 갱신된 GBF/DIME으로 대체하는 수평적 통합(Horizontal Integration)이었다.[11] [그림 6-10]에서 보는 것처럼 대도시 해당지역의 1:100,000 부분을

GBF/DIME으로 대체하여 해당 지역의 타이거를 구축하였다.



출처: Carbauch and Marx, 1990.

[그림 6-10] 지질조사국 지도와 GBF/DIME의 수평적 통합을 통한 타이거 파일의 생성

11) 이 과정은 서로 다른 지도의 도로, 철도, 강 등이 교차하는 지점에서 만나도록 하는 보정 작업을 필요로 한다(Carbauch and Marx, 1990).

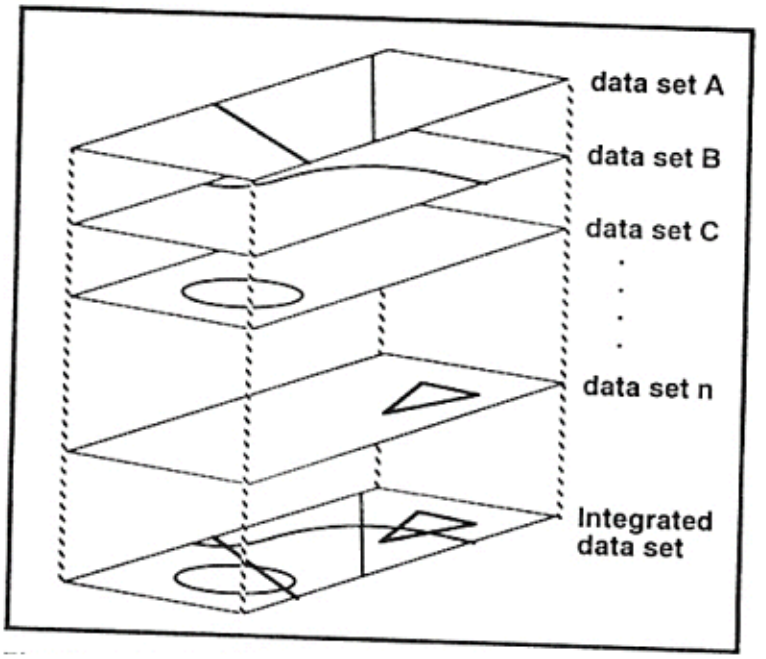
3) 타이거 파일의 사용 및 개선

갱신된 도로 정보, 피쳐 이름, 주소 범위, 지리적 영역의 경계 정보를 포함하는 타이거 파일이 생성되었지만, 검증이 되지는 않았다. 이 예비 파일의 정보를 검증하고 향상시키기 위해 지리실에서는 타이거 파일을 센서스 수행 전에 이루어지는 주소 명부의 편집과 구축 과정에 이용하였다. 대도시 지역의 경우 지리실은 각 주소를 정확한 블록에 할당함으로써 우편조사에 의한 센서스를 수행하기 위하여 센서스국이 취득한 주소 목록과 타이거 파일을 매칭시켰다. 주소와 타이거 파일의 레코드가 일치하지 않으면 타이거 파일을 추가적으로 업데이트하는 것이 필요하다는 것을 의미하였다. 대도시 이외의 지역에 대해서는 현장 조사를 통하여 목록화된 주소를 블록에 할당하였다. 현장 조사원에 의해 지도상의 오류가 발견된 지역과 새롭게 개발된 지역에 대한 지도를 갱신하였다.

나. 타이거 데이터베이스의 구조

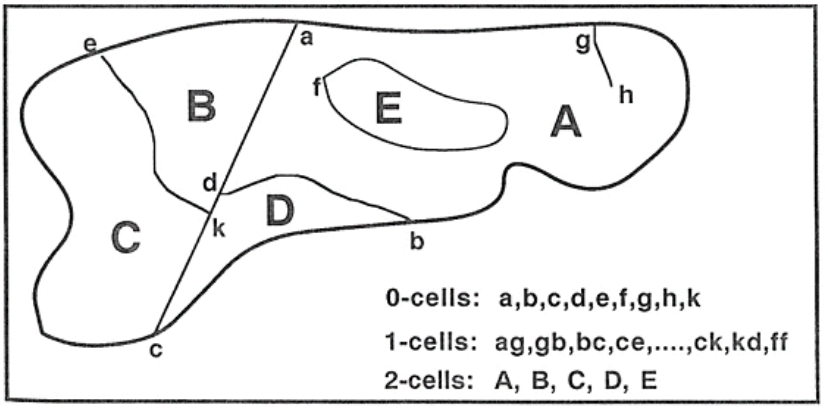
자동화된 지도 생산 시스템으로써의 타이거를 이해하기 위해서는 데이터베이스의 구조에 대한 이해가 선행되어야 한다. 타이거 데이터베이스의 구조를 이해하기 위해서는 개념적 관점과 운영 관점을 구분하여 살펴보는 것이 필요하다. 여기서 개념적 관점은 기본 공간 객체 사이의 관계를 설명하는 것이고, 운영 관점은 파일 관리 시스템 내에서 어떤 파일 구조로 조직화되어 있는지를 설명하는 것이다(Broome and Meixier, 1990).

먼저 개념적 관점은 다양한 개체 유형의 속성 및 관계를 명시한다. 타이거는 다양한 피쳐들을 개별 수준이 아닌 통합된 하나의 수준으로 다룬다는 의미에서 현대적 의미의 GIS에서 일반적으로 사용되는 데이터 구조인 레이어별 데이터 구성과 뚜렷하게 구분된다. [그림 6-11]은 타이거 시스템에서 다양한 공간 데이터 셀들이 하나의 통합된 구조로 구축되는 과정을 개념적으로 표현하고 있다. 이러한 통합 과정에서 발생하는 위상 구조를 명확하게 인코딩하기 위하여 DIME과 유사하게 2, 1, 0 셀의 개념이 사용된다. 즉 실세계의 지리적 현상을 재현하기 위하여 영역, 이들 사이의 경계선 그리고 교차점을 각각 2, 1, 0 셀로 구분한다. 2셀은 더 작은 폴리곤으로 분할될 수 없는 폴리곤을, 1셀은 두 노드를 연결하는 선을, 0셀은 교차점과 끝점과 같은 노드를 나타낸다([그림 6-12]). 1셀은 명시적으로 좌우를 갖고 출발점과 끝점을 연결하고, 2셀과 0셀은 이들과 연결된 1셀에 대한 정보를 갖는다.



출처: Broome and Meixler, 1990.

[그림 6-11] 통합된 하나의 데이터 셀으로 구성된 타이거



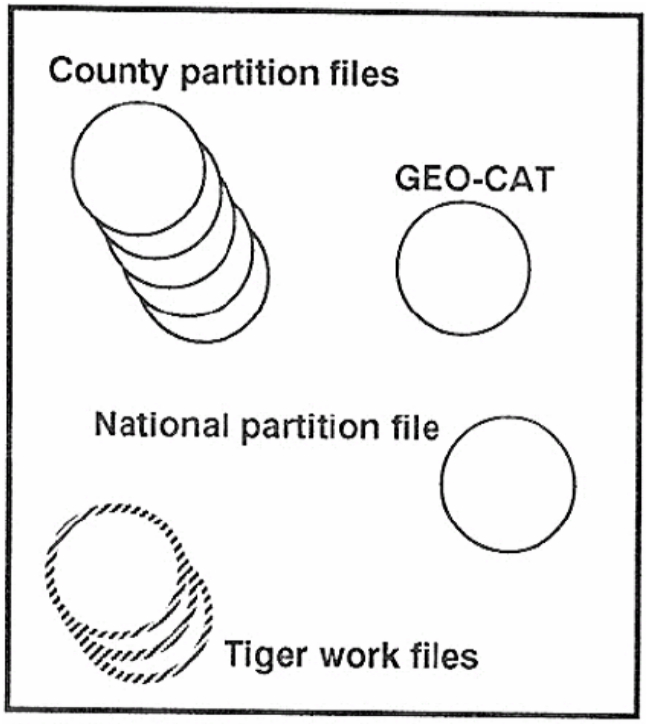
출처: Broome and Meixler, 1990.

[그림 6-12] 타이거에서 0·1·2셀

개념적 관점이 내부 데이터 구조에 초점을 둔다면, 운영 관점은 데이터에 대한 접근

과 조작 방법에 초점을 둔다. 타이거 시스템은 미국 전역을 누락과 중복없이 커버할 수 있도록 구축되었다. 구축된 공간 정보는 블록 단위, 도로망, 하천, 철도, 행정경계, 센서스 경계 등을 포함하는 방대한 데이터베이스로, 물리적으로 하나의 공간에 저장하여 관리하는 것이 불가능하다. 이로 인해 운영 측면에서는 데이터셀을 카운티별로 분할하여 관리한다. 결국 타이거 데이터베이스는 개념적으로는 단일한 데이터베이스이지만, 물리적으로는 독립된 파일들의 집합이다.

운영 관점에서 타이거 데이터베이스는 네 가지 파일 유형 즉, 카운티별 기본 공간 정보를 담고 있는 카운티 분할 파일(County partition files), 행정·통계 단위의 경계에 대한 현재와 과거의 정보를 담고 있는 지리적 카탈로그(GEO-CAT), 카운티 경계 좌표, 개별 분할 카운티 파일에 대한 식별자 정보를 담고 있어서 참조파일 기능을 하는 국가 분할 파일(National partition files), 마지막으로 파일구축 및 갱신과정에서 일시적으로 생성 및 삭제되는 정보를 저장하는 임시 작업 파일(Tiger work files)로 구성된다([그림 6-13]). 타이거 시스템은 카운티에 기반한 약 3,200개의 분리된 파일로 구성되어 있다(Booz and Hamilton, 2000).

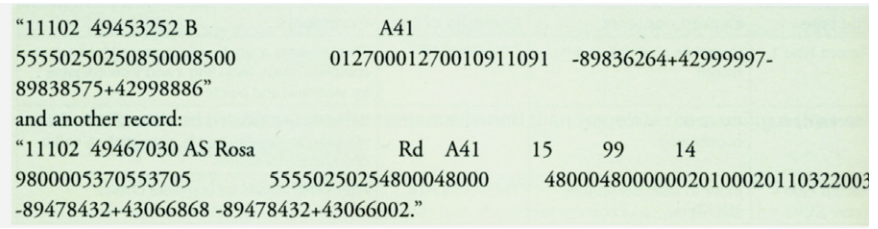


출처: Broome and Meixler, 1990.

[그림 6-13] 운영 관점에서 본 타이거 데이터베이스의 구성

다. TIGER/Line 파일

[그림 6-14]와 같이 텍스트 형식으로 되어있는 타이거 데이터베이스를 이해하고 GIS에서 활용하기 위해서는 TIGER/Line 파일로 변환하는 과정이 필요하다. 이러한 변환 기능을 지원하는 프로그램으로는 ArcGIS의 ArcToolbox에 있는 TIGER-to-coverage Wizard, Maptitude, TransCAD, TGR2SHP, TGR2MIP 등이 있다. 한편 미국 센서스국(http://www.census.gov/geo/www/tiger/index.html)은 카운티 단위로 변환된 TIGER/Line shapefile을 제공한다.



출처: Peters and MacDonald, 2004.

[그림 6-14] 타이거 데이터베이스의 포맷

라. 타이거 데이터베이스의 한계

1990년 센서스를 위하여 구축된 타이거는 이후 갱신 과정을 거쳐 2000년 센서스에서도 사용되었다. 한편 타이거는 현재에도 지속적으로 갱신되고 있다. 국가적 규모의 공간 데이터베이스로서 그리고 센서스의 지리적 업무 지원을 위한 기반으로서 타이거는 오랫동안 기능을 해왔지만 시간이 지나면서 여러가지 측면에서 한계를 드러내고 있다(Trainor, 2003). 타이거 시스템은 1980년대에 센서스국에 의해 개발된 소프트웨어에 기반하고 있는데, 원래 개발되었던 포맷과 운영 환경을 유지하고 있다. 디자인과 플랫폼(flatform)은 개발 당시의 정보기술에 기반하고 있으며, 데이터 관리와 사용자의 접근성에 있어서 현재의 기술적 기대를 충족시키지 못하고 있다. 또한 이 시스템은 실시간의 데이터 업데이트를 지원하지 않는다(Booz and Hamilton, 2000). 타이거 데이터베이스의 한계를 구체적으로 살펴보면 다음과 같다.

먼저 데이터 모델의 관점에서 타이거 데이터베이스는 한계를 지닌다. 위상학적 완결성은 타이거 시스템 요구사항 중 최우선이었는데, 시작 단계부터 타이거는 레이어 기반의 GIS와는 상이한 단일 파일 데이터 구조를 채택하였다. 이 단일 레이어는 기본적인 기하학적 개념에 의해서 이루어진다. 각 끝점, 선, 폴리곤은 명부(list)와 하위 파

일을 통해 연결된 속성 데이터를 갖게 되는데, 이를 통해 지도학자를 포함한 응용프로그램 개발자는 피쳐, 경계, 혹은 영역을 생성할 수 있다. 연결된 도로망을 구축하고 지도학적 기호를 도로에 적용하는 것과 같은 일련의 과정은 별도의 소프트웨어를 통하여 이루어진다. 단일 파일의 위상 기하 구조는 현재 널리 사용되고 있는 피쳐-오브젝트(feature-object) 구조 이전의 것이다. 일반 사용자의 경우, 상업용 GIS로 데이터를 변환하기 위하여 TIGER/Line 파일 변환기가 필수적이다. 센서스국 직원의 경우, 내부에서 개발된 소프트웨어를 이용하여 데이터를 읽고 지도를 작성한다. 이러한 단일 레이어 위상 구조 데이터베이스는 타이거 데이터에 대한 활용을 제약한다. 센서스국 내부에서조차 복잡한 내부 시스템의 도움 없이 주제도를 작성하는 것이 어려웠다. 이에 따라 1990년대 초에 센서스국의 데이터 공표 활동을 지원하는 지도학적 기능은 상업적 GIS를 활용하는 것으로 변화하였다.

타이거에서는 새로운 데이터 레이어를 추가하는 것이 어려운데, 이는 엄격하고 제한된 데이터 구조에 기인한다. 20여 년이 지난 시점에서 볼 때, 타이거는 기술변화를 유연하게 받아들일 수 없는 데이터 구조를 가지고 있다. 또한 타이거 데이터베이스를 구축하는데 사용된 원자료의 정확도와 품질이 균질적이지 않다는 점도 타이거를 지속적으로 유지하는데 부담으로 작용한다. 타이거 데이터베이스를 구성하는 공간 데이터와 그 속성은 GBF/DIME 파일, 지질조사국의 1:100,000 지형도, 지방 지도 등 다양한 자료원에서 추출되었다. 기존 데이터의 편집을 통하여 지도 베이스(map base)를 형성하였는데, 여기에 도로·경계와 같은 데이터들은 현장 조사 활동을 포함하는 센서스국의 자체 프로그램을 반영한 결과로 만들어졌다. 상이한 자료원과 작업 방식은 타이거 데이터베이스의 질에 직접적인 영향을 주었다. 몇몇 연구들에서 타이거 파일의 정확도에 대한 문제제기와 평가가 이루어졌다. Ratcliffe(2001)는 지적(cadastral) 자료와의 비교를 통하여 고해상도의 지오코딩된 데이터에서 개별 주소의 위치 정확도를 평가하였다. 그 결과 5～7%의 주소가 잘못된 센서스 트랙에 할당되었다. 토지 필지 수준에서는 50% 이상이 일치하지 않았다. Liadis(2000)에 따르면, 센서스국의 지리실은 타이거 현대화를 위한 준비의 일환으로 타이거 데이터베이스의 공간적 정확도를 평가하기 위하여 GPS(Global Positioning Systems)를 이용할 계획이다. GPS를 이용하여 수집한 포인트와 이에 상응하는 타이거 0 셀 사이의 거리를 비교한 결과, 타이거와 실제 측정 지점 사이의 평균 거리 차이에서 상당한 변이가 존재하였다. 한편 O'Grady(2001)는 DOQ(Digital Orthophoto Quadrangles)와 GPS를 이용하여 타이거를 향상시키는 방안을 논의하였다. DOQ를 바탕으로 새로운 피쳐들을 추가하고 기존 피쳐들을 공간적으로 강화하며, GPS 기술을 이용하여 갱신된 타이거 데이터베이스를 검증하는 방안을 제시하였다. 이러한 타이거의 위치 정확도 향상은 센서스국 내부적으로는 새로운 좌표 취득을 위한 강력한 방법으로 GPS 사용을 고려

하면서 시작되었으며, 외부적으로는 중앙 및 지방 정부 사이의 데이터 교환을 용이하게 할 필요에 의해서 유발되었다(O'Grady and Godwin, 2000).

초기에 센서스국은 타이거 시스템의 잠재력을 완전하게 인식하지 못하였다. 그들의 초점은 센서스를 지원하기 위한 시스템의 유지·보수, 변화에 따른 데이터의 갱신, 법적·행정적 경계의 위치 갱신 등이었다. 타이거 데이터를 이용하기 위해서는 상용 프로그램을 부분적으로 모방한 복잡한 소프트웨어 개발을 필요로 하였다. 직원들, 특히 관리자의 GIS에 대한 제한적인 지식과 함께 값비싼 소프트웨어 대한 제한적인 접근은 그 사용을 제약하였다.

마. MAF/TIGER 강화 프로그램

앞에서 언급한 타이거 데이터베이스의 한계, 특히 위치 정확도 문제를 해결하기 위해 센서스국은 2002년부터 MAF/TIGER 강화 프로그램을 시작하였다.[12] 미국의 센서스는 우편 조사에 기반하고 있기 때문에 정확한 주소 명부가 센서스에 있어 필수적이다. 1990년 센서스 이후 센서스국은 처음으로 2000년 센서스를 위하여 MAF라는 영구 주택단위 주소 명부 개발에 착수했다. 이 전에는 10년 주기 센서스 때마다 명부가 작성된 후 폐기되었다. 이는 1990년 구축된 가구 단위에 대한 Address Control File(ACF)과 USPS의 DSF(Delivery Sequence File), PALS(Programs for Address List Supplement)를 통하여 취득된 지방의 주소 목록, DSF를 사용하지 않는 지역에 대한 주소 조사 목록을 바탕으로 구축되었다.[13] MAF/TIGER 강화 프로그램의 중요 핵심 목표는 다음과 같다(Trainor, 2003).

□ 도로, 기타의 피쳐 및 거처에 대한 정확한 위치 제공, 자동화된 피쳐-변화 감지 기법을 이용한 새로운 도로, 거처 단위 확인
□ 상업용 소프트웨어와 GIS 툴에 기반한 새로운 MAF/TIGER 처리 환경의 개발
□ 지리적 파트너십 프로그램의 확장 및 장려와 커뮤니티 주소 갱신 시스템(Community Address Updating System: CAUP) 착수

이러한 프로그램을 통하여 궁극적으로 추구하고자 하는 것은 공간 데이터에 대한 상대적 정확도가 아닌 절대적 위치 정확도를 향상시키는 것이다. 1990·2000년 센서스 지원을 목적으로 구축된 타이거 시스템은 그 목적을 성공적으로 달성하였고, 타이거 파일은 여전히 국가적 자원으로 기능을 하고 있다. 그러나 기술 발전은 타이거의 상대적 정

12) MAF는 Master Adress File의 약자이다.
13) USPS의 DSF는 주소들을 확인해주고, 누락 주소에 대한 정보를 준다.

확도를 유지하는 센서스국의 능력을 초월하고 있다. 조사원이 거처를 정확하게 위치시키는데 도움을 주는 모바일 컴퓨팅 기기들은 이전과는 상이한 수준의 정밀도(precision)와 정확도(accuracy)를 요구하고 있고, 결과적으로 타이거 데이터의 질에 직접적으로 영향을 준다. 모바일 컴퓨팅 기기의 활용은 현재 타이거에서 이용가능한 것보다 더 높은 위치 정확도를 요구한다. 따라서 타이거 데이터의 위치를 향상시키는 것, 그리고 좀 더 일관되게 정확한 데이터를 취득하는 것이 MAF/TIGER 강화 프로그램의 첫 번째 목표이다. 상대적 위치 정확도에서 절대적 위치 정확도로의 시스템의 목표 변화는 데이터 구축에서 좀 더 엄격한 규칙의 적용이 필요하고, 고품질의 데이터 표준이 요구된다는 것을 의미한다.

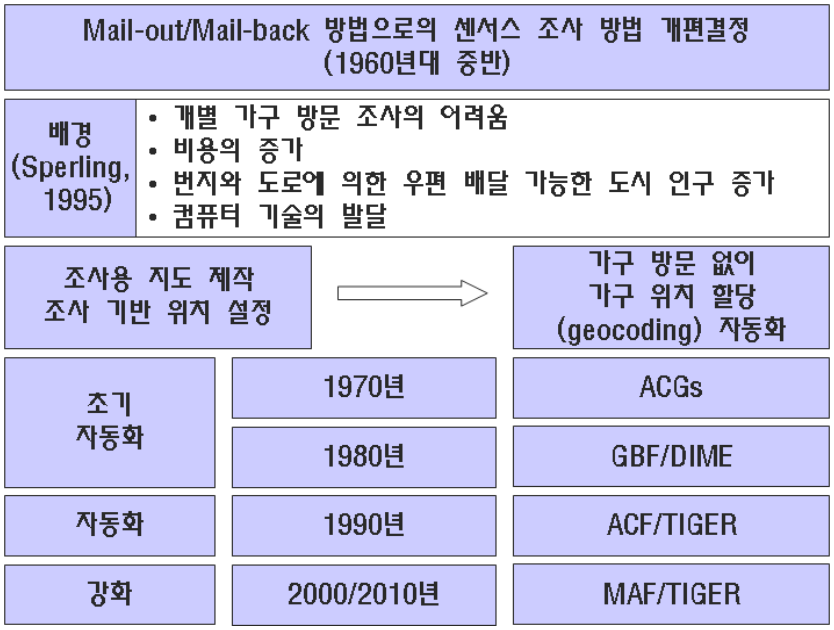
두 번째 목표는 상업용 소프트웨어와 GIS 툴에 기반한 새로운 MAF/TIGER 처리 환경의 개발이다. 상업용 GIS에 의해서 타이거 데이터를 읽게 된 것은 최근의 일이다. 원래 이 시스템은 상업용 소프트웨어의 도움없이 만들어졌다. 당시 GIS산업은 발전의 초기 단계였으며, 전산화된 공간 정보의 생성은 높은 비용에 의해 주로 제한되었다(Trainor, 2003). 모든 측면에서 이 시스템은 내부에서 계획되고, 개발되고, 프로그램화 되었다. 데이터 모델, 데이터 구조, 데이터 입력기능, 응용(지도화, 지오코딩, 업데이트), 산출물의 생성은 센서스국의 직원에 의해 계획되고, 개발되고, 작동되었다.[14] 과거에는 담당 직원에 의해 파생 데이터베이스가 생성된 후 변환되고 그런 다음 그것을 공간 데이터베이스에 올려 놓았다. 이러한 변환 과정은 상당한 시간과 노력을 소요하는 작업이지만 GIS 응용 소프트웨어의 활용을 위해서는 필수적인 단계였다. 그 다음 담당 프로그래머가 상용 GIS에서 읽을 수 있는 형태로 타이거 데이터를 추출할 수 있는 프로그램을 개발한다. 현재 타이거 시스템은 센서스국의 필요와 요구에 집중해 있는 상태지만, 미래의 강화된 타이거는 국가 공간 데이터 인프라에 대한 투입물을 제공할 것이다. 타이거의 일차적인 목적은 센서스의 지원이지만, 향후 수많은 잠재적 사용자의 요구를 만족시키기 위한 고려가 이루어지고 있다. 타이거의 광범위하고 다양한 사용은 그것을 만든 사람들을 놀라게 했다. 사용자들로부터의 반응과 제안은 미래의 타이거를 논의하고 계획하는데 영향을 미친다.

세 번째 목표는 데이터 구축에서 기존의 연방정부 중심의 하향식 접근에서 지방 정부와의 파트너십을 통한 상향식 접근으로의 데이터 구축 방법론의 변화이다. 이러한 변화는 분권화된 환경에서 데이터베이스의 효과적인 갱신을 보장할 것이다. 2000년 센서

14) 프로그래밍의 특성은 이야기를 쓰는 것과 유사하다. 작가나 프로그래머는 그들 자신의 스타일과 방식을 가지고 있다. 시간이 흐름에 따라 소프트웨어가 발전하고, 환경이 변화함에 따라 전산 코드를 유지하는 것은 도전에 직면하게 된다. 종종 다른 프로그래머의 코드를 고치는 것보다 다시 쓰는 것이 쉽고 효과적이기도 하다. 독립적으로 개발된 프로그래밍 스타일은 공급자나 사용자의 상호작용을 제한하는 경향이 있다.

스에서 주소 명부의 질은 LUCA(Local Update of Census Addresses) 프로그램에 달려 있었다. 센서스국은 지방정부와 도로 주소 데이터베이스의 업데이트를 위해 지방정부와 공동작업을 수행해 왔다. 도시와 카운티는 도로 건설과 주소를 지정하는 과정을 통제하기 때문에 이에 관한 최상의 정보를 가지고 있다. 경우에 따라 조사원이 직접 주소 파일에 없는 거처를 현장 방문해서 확인하는 경우도 있었다(Peters and MacDonald, 2004).

이상에서 살펴본 1970년대 이후 센서스 방법론의 변화와 지리적 지원 기능의 변화를 도식화해서 살펴본 결과는 다음과 같다([그림 6-15]).



[그림 6-15] 1970년대 이후 센서스 방법론의 변화와 지리적 지원 기능의 변화

제3절 조사 기반 센서스 실시 국가에서의 GIS

통계지리시스템은 센서스 조사의 계획, 데이터 수집, 처리 및 유포 등 센서스 전 과정에서 이용될 수 있다(UN, 2008). 다음에서는 센서스 단계별로 GIS가 지원할 수 있는 역할에 대해 구체적으로 살펴본 후 주요 국가의 센서스 GIS에 대해 살펴보았다. .

1. 센서스 단계별 GIS의 역할

가. 센서스 전단계(pre-census)

1) 표준 통계 단위의 설정과 관리

GIS의 지리정보 및 디지털 지도는 센서스가 실시되기 전 각종 통계 단위의 설정과 체계적 관리를 촉진시키고 일관성을 유지하게 한다. 통계 단위의 경계에 따라 표본 추출이나 결과가 영향을 받기 때문에 통계 수집의 기본이 되는 공간 단위의 획정과 관리는 매우 중요하다(Gould and Hecht, 2001). 각 나라마다 상황이 다르기 때문에 서로 다른 공간 단위를 획정하여 사용하고 있지만, 일반적으로 센서스에서 사용되는 공간 단위는 크게 법적·행정적 단위(legal and administrative entity)와 통계 데이터 수집과 공표 목적의 통계 단위(statistical entity)로 구분될 수 있다.[15] 행정 단위는 중앙 정부 및 지방정부, 유관 기관들의 행정·관리를 위한 경계이며, 일반적으로 계층적 위계를 가진다.[16] 통계의 집계는 일반적으로 이러한 행정 단위에 따르기도 하지만, 다양한 수요를 만족시키기 위해 이차적인 지리적 범주를 통해 재구획화 되기도 한다. 통계 단위는 데이터 사용자의 실용적 또는 분석적 요구에 부합할 수 있는 지리적 범위를 토대로 개발되기 때문에, 특정 계층 수준에 있어서는 행정 단위와 완전한 포함관계를 가지지 않고, 상이한 구획을 보이기도 한다. 통계 단위 역시 일반적으로 계층적 위계를 가지며, 우편구역, 도시지역 등의 기능지역들이 단독적으로 존재하기도 한다. 대부분의 국가 통계 기관에서는 센서스 수행에 있어 행정 단위와 통계 단위가 혼용되어 사용되고 있다.

이러한 통계 단위 중 가장 유명한 것 중 하나는 미국 센서스국의 블록, 센서스 트랙 등인데, 미국의 1990년 센서스의 공간 단위를 살펴본 결과는 다음과 같다(<표 6-2>). 통계 단위는 종종 실용성, 관행, 필요에 의해 만들어졌지만 미국 센서스국은 그것의 식별과 경계구분을 위한 기준(criteria)과 가이드 라인을 점차 발전시켰다. 법적·행정 단위는 그 경계가 공식적으로 주어져 있는데 반해, 통계 단위는 법적·행정 단위만으로는 포괄할 수 있는 범위가 제한적이거나 부적절하거나, 시계열적으로 일관되지 않거나 혹은 존재하지 않는 상황에 적절하다. 어떤 사용자는 시계열 비교를 허용하는 안정적인 경계를 필요로 하는 반면 어떤 사람들은 변화하는 상황을 반영하는 업데이트된 경계를 요구하기도 하는데, 센서스국은 통계 단위를 데이터 사용자의 실용적인 측면과 분석적인 측면에서의 요구에 부응하기 위해 발전시키고 있다. 미국의 2000년 센서스에서 사전에 조사

15) 우리나라의 도-시군구-읍면동, 미국의 State-County-City-Incorporated places 등이 행정 단위의 예이며, 우리나라의 조사구-집계구, 미국의 census block-census block groups-census tract 등이 통계 단위의 예이다

16) 미국에서 학군은 다른 공간 단위에 비해 새롭게 등장한 법적·행정적 단위인데, 1980년과 1990년에 센서스국에서는 학군에 기초한 사회경제적 정보를 제공하였다. 이 경우 센서스국은 교육부와 각 주의 담당 부서와 전국의 학군의 지리적 범위를 정하기 위해 협력하였다.

한 연방 기관별 요구 공간 단위를 살펴보면 법적·행정적 단위뿐 아니라 통계 단위에 대한 요구가 많음을 알 수 있다(<표 6-3>).

〈표 6-2〉 미국 1990년 센서스의 공간 단위

공간 단위 유형	구분	숫자
국가	**법적 단위**	**1**
지역(Regions)	**통계 단위**	**4**
지구(Divisions)	**통계 단위**	**9**
주(States)	**법적 단위**	**57**
카운티(Counties)	**법적 단위**	**3,248**
하위 카운티와 장소(County Subdivisions and Places)		**60,228**
Minor Civil Divisions(MCDs)	법적 단위	30,386
Sub-MCDs	법적 단위	145
Census Counry Divisions	통계 단위	5,581
Unorganized Territories(UTs)	통계 단위	282
Other Statistically Equivalent Entities	통계 단위	40
Incorporated Places	법적 단위	19,365
Consolidated Cities	법적 단위	6
Census Designated Places(CDPs)	통계 단위	4,423
인디언과 알래스카 원주민 보호구역		**576**
American Indian Reservations(no trust lands)	법적 단위	259
American Indian Entities with Trust Lands	법적 단위	52
Tribal Juridiction Statistical Areas(TJSAs)	통계 단위	19
Tribal Designated Statistical Areas(TDSAs)	통계 단위	17
Alaska Native Village Statistical Areas(ANVSAs)	통계 단위	217
Alaska Native Regional Corporations(ANRCs)	법적 단위	12
대도시권(MAs)		259
Metropolitan Statistical Areas(MSAs)	통계 단위	362
Consolidated Metropolitan Statistical Areas(CMSAs)	통계 단위	268
Primary Metropolitan Statistical Areas(PMSAs)	통계 단위	21
도시화 지역(Urbanized Areas: UAs)	**통계 단위**	**73**
특수 목적 지역		**404,583**
의회선거구	법적 단위	435
투표지역	법적 단위	148,872
학군	행정 단위	15,274
교통분석지역	행정 단위	200,000
ZIP코드	행정 단위	40,000
센서스 트랙과 BNBs(Block Numbering Areas)		**62,276**
Census Tracts	통계 단위	50,690
Block Numbering Areas	통계 단위	11,586
블록 그룹(Block Groups)	통계 단위	**229,192**
블록(Blocks)	통계 단위	**7,017,427**

www.census.gov/geo/www/GARM/Ch2GARM.pdf.

〈표 6-3〉 미국 2000년 센서스에 대한 연방 기관별 요구 공간 단위

구분	공간 단위
농무부	카운티
	센서스 트랙
	인디언과 알래스카 원주민 보호구역
상무부	국가
	카운티
	Place
	학군
	센서스 블록
	인디언과 알래스카 원주민 보호구역
국방부	Zip 코드
	센서스 블록
교육부	학군
	센서스 블록
보건부	카운티
	Place
	Zip 코드
	인디언과 알래스카 원주민 보호구역
주택・도시개발부	Place
	센서스 트랙
	블록 그룹
법무부	Place
	학군
	센서스 블록
노동부	카운티
	센서스 트랙
교통부	교통분석지역
	센서스 블록

출처: www.census.gov/geo/www/GARM/Ch2GARM.pdf.

2) 센서스 GIS 데이터베이스 구축

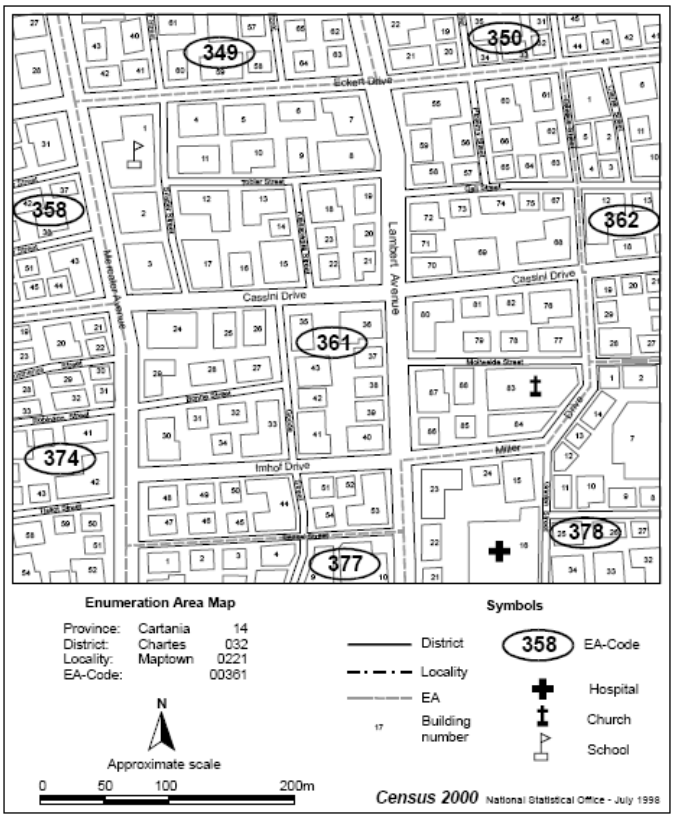
조사구(enumeration area)는 조사원들에게 할당되는, 센서스 데이터 수집을 위한 기본

적인 공간 단위이며, 센서스의 초기 단계에 정의된다. 조사구는 일반적으로 통계 단위 중 가장 최소 단위로 생성되기 때문에 보다 상위에 있는 통계 단위의 생성을 위한 기반이 된다. 조사구의 획정에 필요한 다양한 기준들은 다음과 같다(UN, 2008).

① 중복되지 않게 상호 배타적이어야 하며, 국가 전체를 포괄해야 한다.
② 쉽게 확인될 수 있는 경계를 가져야 한다.
③ 정부 부처나 다른 데이터 사용자들의 요구들을 충족시켜야 한다.
④ 행정적 계층과 일관성을 가져야 한다.
⑤ 다른 센서스와 데이터 수집 활동에 유용해야 한다.
⑥ 불연속이나 빈 틈(hole)이 없이 치밀해야 한다.
⑦ 유사한 인구 규모를 가져야 한다.
⑧ 센서스 기간 동안 조사원에 의해 포괄(cover)될 수 있게 작고, 접근 가능해야 한다.
⑨ 상이한 통계 공표 단위별로 광범위한 테이블을 만들 수 있도록 작고 유연해야 한다.
⑩ 데이터의 개인정보를 보호할 수 있도록 충분히 커야 한다.

이러한 기준들 중 일부는 데이터의 수집과 관련되며, 일부는 집계 결과 생산과 관련된다. 센서스의 목적이 유용한 통계 데이터를 생산하는 것임을 감안할 때 조사구의 유연성과 적합성이 조사의 편의성보다 우선시 되어야 할 것이다(UN, 2008). 그럼에도 불구하고, 조사원에 대한 적절한 업무량 할당이라는 현실 문제의 고려 또한 간과되기는 힘들다. 조사구는 현장 조사원의 실제적인 운영에 있어 필수적인 부분이기 때문에, 조사원이 분배된 시간에 하나의 조사지역을 완전히 조사할 수 있도록 디자인 하는 것이 중요하다. 따라서 조사원의 적절한 작업량과 조사 시 발생할 수 있는 물리적 제약들, 예를 들어, 낮은 접근성을 고려한 기준 등을 제시하고, 이러한 기준들을 충족시키는 적절한 크기의 조사구를 설정하는 것이 요구된다.

GIS는 센서스와 결합되어 위와 같은 기대와 기준들을 만족시킬 수 있는 조사구 설정을 위한 유연한 기능들을 제공할 수 있다. 이와 함께 조사원에게 할당된 지역의 경계, 학교・도로・병원 등과 같은 이동 방향에 대한 가이드를 제시해 줄 수 있는 지리적 객체들이 표시되어 있는 조사구요도를 제공할 수 있다([그림 6-16]). GIS를 이용하여 제작된 디지털화된 조사구요도는 지도의 크기・심벌・컬러・스케일・레이블 조정 등의 지도학적 요소들을 손쉽게 변경할 수 있으며, 현장에서 요구되는 수만큼 신속하게 제작될 수 있다.



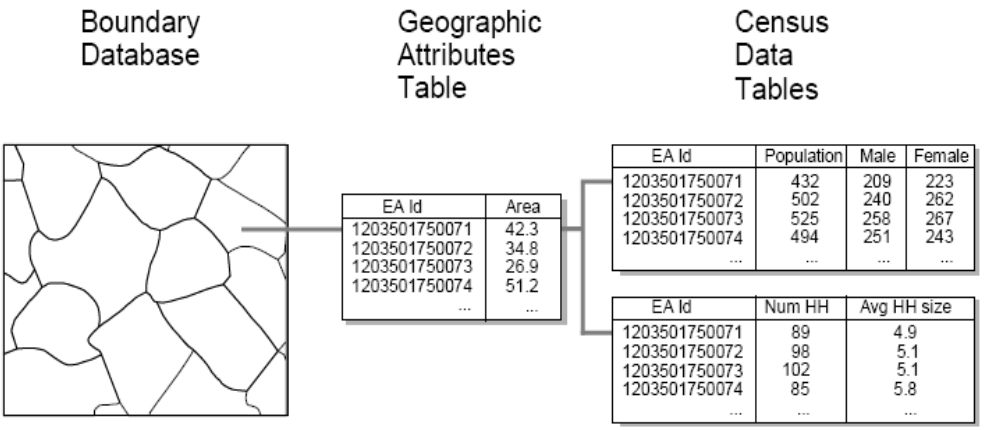
출처: UN, 2008.

[그림 6-16] 조사구요도

센서스와 공간 데이터베이스의 결합은 센서스 속성들을 실세계의 위치들과 연결하는 지오코딩을 통해서 이루어지진다([그림 6-17]). 센서스 속성과 실세계 위치의 연결은 개별 주택 단위에서부터 조사구 및 상위의 공간 단위를 모두 포괄한다. 센서스 조사표의 주소는 GIS를 이용하여 필지나 건물의 공간적 객체와 연결되며, 조사구를 포함하는 각종 통계 단위로 합역된(aggregated) 센서스 속성들은 고유한 식별자를 통해 공간 데이터베이스 내에 있는 디지털 구역들과 연계된다. 각 통계 단위 역시 계층적 위계에 따라 고유 식별자를 통해 연계된다.

센서스 공간 데이터베이스는 통계 단위가 전통적인 경계를 벗어날 수 있는 가능성을 제공하며, 요구에 따라 데이터를 다양한 형태로 합역할 수 있도록 해준다. 또한 연계된 속성 데이터를 지리적 형상으로 시각화함에 따라 데이터 결측을 쉽게 확인할 수 있으며, 공간 내삽(spatial interpolation)과 같은 GIS 기술을 통해 결측 데이터를 추정 보완할 수

있다. 현장 조사를 위해 모집단으로부터 표본을 추출하기 위한 단계에서도 센서스 공간 데이터베이스는 유용하게 이용될 수 있다(Amar and Shahar, 2001).



출처: UN, 2008.

[그림 6-17] 지오코딩을 통한 공간 데이터와 속성 데이터의 연계

나. 센서스 단계(census)

센서스에서 디지털 맵핑의 주요한 기여는 주로 조사 이전과 이후이지만, 실제 조사 단계에서도 다음과 같은 여러가지 역할을 할 수 있다.

1) 현장 조사의 관리와 제어 및 업데이트

GIS는 현장 조사를 직접 수행하는 조사원들의 효율적 활동을 위한 여러 가지 운용 계획이나 조사구요도와 같은 실질적인 자료를 제공하는 한편, 관리자에게도 유용한 도구를 제공할 수 있다. GIS는 자동화된 프로세스를 통해서 관리를 위한 지도를 생성할 수 있고, 디지털화된 주소 파일을 제공할 수 있어, 조사원이 수행하는 업무의 질적·양적 제어를 손쉽게 할 수 있게 해준다. 한편 실제 현장 조사 시 발견되는 새로운 건물이나 변경 사항 및 에러들을 실시간으로 공간 데이터베이스에 입력하여 업데이트하여, 이후 통계 생산을 신속하게 진행할 수 있도록 해준다.

2) 센서스 과정의 모니터링

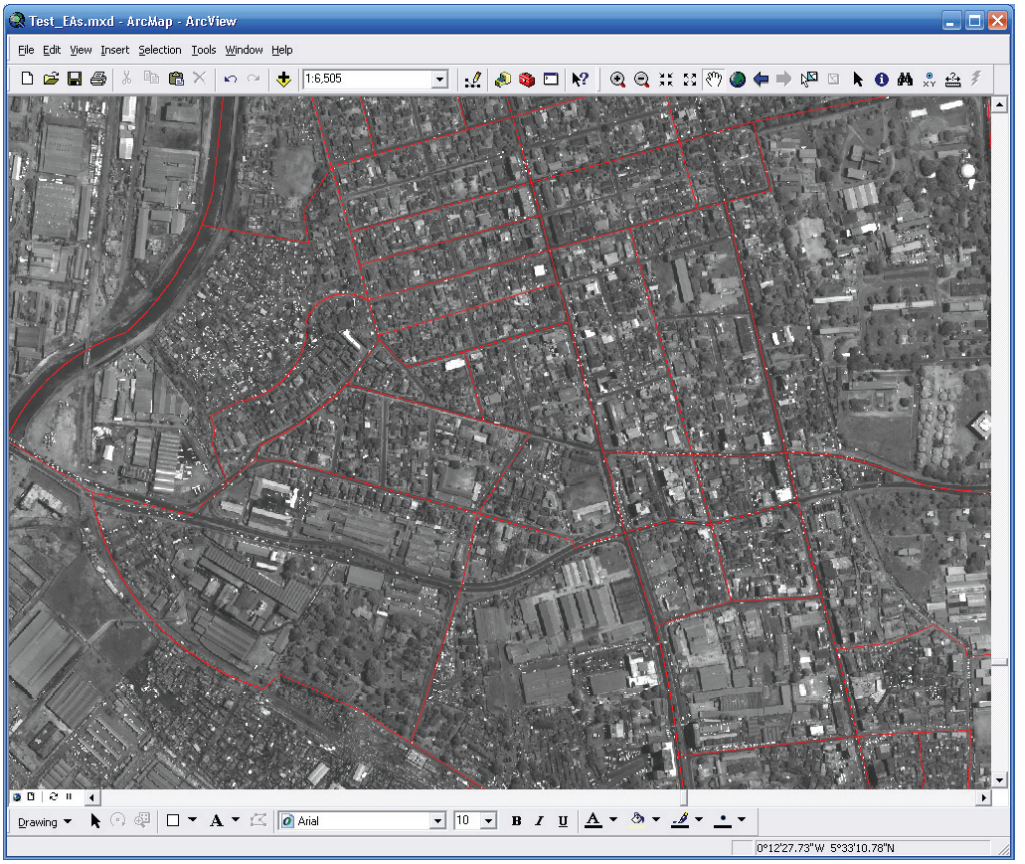
조사 기간 동안, 센서스의 본부 관계자들은 조사와 데이터 가공의 진척 상황을 모니

터링할 것이다. 지방 사무소가 가장 먼저 조사 활동과 결과에 대한 정보를 모으고, 본부에서는 이 정보를 수집하여 어느 곳의 운용이 원활하게 되었는지, 어떤 문제가 발생했는지를 평가할 것이다. GIS를 통해 진척 정도가 현저하게 낮거나 높은 지역들을 시각적으로 신속히 확인할 수 있으며, 각 지역의 조사 진척이나 완료 상태를 전반적으로 평가할 수 있는 지도 역시 제작될 수 있다.

센서스의 참여에 대한 관심과 응답률의 하락은 성공적인 센서스 수행에 있어 걸림돌이 될 수 있다. 센서스에 대한 관심과 참여를 고취시키기 위한 노력은 중앙 정부뿐 아니라, 센서스 수행에 있어 가장 현장 가까운 곳에 위치하는 지방정부의 역할 또한 중요할 것이다. GIS는 여러 기관들이 센서스 참여를 촉진시키기 위해 할 수 있는 여러 활동들을 체계적으로 지원할 수 있다. 예를 들어, 최근의 전출입자, 외국인, 저소득 계층 등 자료 수집이 어려운 인구 부문이 많은 주요 지역 파악, 비응답 인구의 지도화, 센서스 데이터의 유용성에 대한 지역 주민과의 의사소통, 지역 보도 자료를 위한 자료 생성 등에 GIS를 이용할 수 있다(Beddoe, 1999).

3) 신기술과의 결합을 통한 새로운 조사 방법의 도입

고해상도 위성 이미지, GPS, PDA(Personal Digital Assitant, 개인휴대정보단말기)와 같은 모바일 기기의 등장은 기존 GIS의 응용 영역을 한층 확대시켰다. 최근까지 도시 지역의 지리 데이터 획득은 항공사진이 필수적이었고, 인공위성 이미지는 저해상도로, 범람원이나, 도시 오염, 토지이용과 같은 연구에 제한되었다. 하지만 최근의 고해상도 이미지는 높은 비용에도 불구하고, 위치 정확성, GIS와의 연결성, 잦은 업데이트 주기 등으로 인해 이용이 증대되고 있다(United Nations, 2004). 고해상도의 인공위성 이미지는 센서스 조사에도 유용하게 이용될 수 있는데, 접근이 어려운 지역에 대한 인구 추정, 도시화된 토지이용의 측정, 거주 인구 밀도의 추정을 위한 토지피복 변화의 확인 등에 이용될 수 있고, 나아가 통계 단위의 재조정을 위한 자료로 이용될 수도 있다. 또한 과거 조사구의 토지 이용과 지형 변화 등을 쉽게 확인할 수 있어, 별도의 주의가 요구되는 지역을 신속하게 파악할 수 있게 해줌으로써, 이후 데이터베이스의 업데이트나 센서스 계획을 위한 현실성 있고, 상세한 근거가 될 수 있다. 직접적인 답사 없이 현장에 대한 상세한 지리적 정보를 제공할 수 있기 때문에, 현장 조사원의 조사 계획이나 스케줄링에도 유용하게 이용될 수 있으며, 디지털 조사구 경계와 함께 위성 이미지나 또는 정사사진(orthophotograph)을 중첩하여 보다 현실성 있는 조사구요도를 조사원에게 제공할 수 있다([그림 6-18]). 벨기에 통계청에서는 위성 이미지를 사용하고 있으며(Jamagne, 1998), 우리나라 통계청에서도 2009년 11월에 실시한 2010년 인구주택총조사 시범예행조사에서 항공사진을 결합한 조사구요도를 시험 서비스하였다.



출처: UN, 2008.

[그림 6-18] 인공위성 이미지를 결합한 조사구요도

GPS 기술은 지표상의 정확한 위치 확인을 가능하게 해주기 때문에, 조사구의 정확한 경계 설정과 현장의 정확한 지리적 객체를 저비용으로 효율적으로 파악할 수 있도록 지원한다. GPS 기술과 더불어, PDA와 같은 모바일 기기의 등장은 센서스 데이터 수집 및 배포에 획기적인 역할을 할 수 있다. 현재 많은 GIS 회사들은 제한된 기능을 가지지만 PDA나 태블릿 컴퓨터와 같은 모바일 장비에서 운용될 수 있는 소프트웨어를 제공하고 있다. 이러한 장치들을 이용한 휴대용(hand-held) GIS는 미국 센서스국, 농무부 농업통계국(National Agricultural Statistics Service)과 같은 통계 기관에서 현장 데이터를 수집하는데 널리 사용되고 있다. 이러한 통계 기관에서는 현장에서 데이터를 저장하고 사무실로 전송할 수 있도록 현장 조사원이 일상적으로 PDA를 휴대하게 하고 있다(이상일 외, 2009). 현장 조사원은 GPS 수신기가 장착된 PDA를 휴대하여, 조사

된 데이터 및 운용 시에 발생하는 제반 상황들을 실시간으로 저장하고 보고할 수 있다. 미국 센서스국은 FDCA(Field Data Collection Automation)라는 프로그램을 통해, 2010년 센서스 조사에 조사표가 취합되지 않은 세대들의 재방문 지원 계획을 수립했다. 조사원들은 GPS 수신이 가능하고, ESRI社의 ArcPad와 같은 모바일 GIS 프로그램을 탑재한 모바일 기기를 휴대한다. 이를 통해 조사원들은 무응답 주소의 정확한 GPS 위치와 방문 경로를 신속히 확인하고, 데이터베이스와의 직접적 연결을 통해 작업 운용의 효율성을 증대시킬 수 있을 것이라 예상되고 있다. 미국의 2010년 센서스의 현장 조사 기반은 FDCA를 운용할 13개 지방청과 455개의 지역 사무소로 구성될 것이고, 약 50여만 명의 조사원이 데이터 수집 작업에 동원될 것이다(ESRI, 2006). GIS와 모바일 기기의 결합은 센서스 결과의 배포에도 유용하게 이용될 수 있다. 무선 인터넷 기술은 데이터를 각 가정이나 사무실에 설치된 PC뿐 아니라, 이동 중인 개인, 차량 등에 탑재된 모바일 터미널로 신속히 전송할 수 있게 해준다. 다양한 형태의 새로운 기술들은 GIS 시스템과 보다 긴밀하게 연결되어 센서스 작업들을 보다 효율적이고, 진보된 형태로 수행할 수 있게 도울 것이다.

다. 센서스 사후단계(post-census)

센서스 결과의 집계와 집계된 통계 및 각종 분석 결과를 대중에게 유포하는 이 단계에서는 사회 각계 각층의 수요에 대해 능동적으로 대처할 수 있는 유연성이 무엇보다 중요한데, GIS는 이의 중요한 기반이 될 수 있다.

1) 자료 집계

센서스 조사 후 가장 중요한 작업 중 하나는 취합된 데이터를 적절한 통계 단위로 집계하는 것이다. 정부 기관, 연구원, 기업, 일반 사용자 등 매우 다양한 영역의 센서스 데이터 수요자들은 그들의 목적에 따라 상이한 공간 단위로 데이터가 제공되기를 원한다. 예를 들어, 행정 목적에 필요한 통계 데이터를 원하는 지방정부는 행정 단위에 따라 집계된 데이터가 필요할 것이고, 시장 점유율이나 고객 마케팅을 위해 데이터를 필요로 하는 기업은, 행정 단위보다 상세한 수준의 집계 자료를 원할 것이다. 따라서 이러한 다양한 수준의 수요를 만족시키기 위해서는 최소 공간 단위인, 조사구에 따라서 취득된 데이터들이 다양한 지리적 수준으로 합역될 필요가 있다. 이러한 합역 작업은 공간 데이터베이스 내에서 쉽게 수행될 수 있으며, 나아가 단순한 합역으로 생성될 수 없는 공간 단위의 통계 역시, 공간 내삽이나 위상학적 공간관계에 기반한 GIS의 기능을 이용하여 효과적으로 생산될 수 있다.

2) 센서스 결과의 주제도화

데이터가 수집되고 나면 통계청은 다양한 형태의 센서스 결과를 일반 대중을 포함한 다양한 수요자들에게 효과적으로 유포하기 위한 방법을 모색할 것이다. 통계 데이터의 전달과 해석에 있어 단순한 숫자나 텍스트의 테이블보다 그래픽이나 시각적 도구를 이용하는 것이 효과적이라는 것은 이론의 여지가 없다. 특히 지역을 기반으로 하는 통계 결과는 지도화를 통해서 보다 그 유용성을 확대시킬 수 있다. 각종 통계 자료의 주제도화 및 지도화는 다음과 같은 점에서 통계 결과를 표현하는데 효과적이다 (UN, 2008).

① 지도는 전달하고자 하는 아이디어나 개념을 효과적으로 전달할 수 있다.
② 말로 설명하기 어려운 주제들이나 이슈들이 지도나 그래픽 형태로 효과적으로 묘사될 수 있다.
③ 지도는 보는 사람의 호기심을 자극한다. 독자의 주의를 끌게 하고 부가적인 정보를 담고 있는 텍스트를 함께 읽도록 촉진한다.
④ 지도는 많은 정보를 간략하게 요약해준다. 많은 양의 숫자들이나 관측치 사이의 공간 관계를 효과적으로 전달할 수 있다. 또한 데이터의 높고 낮은 값들의 군집 역시 지도가 아닌 일반 데이터 표현으로는 파악하기 어렵다.
⑤ 지도는 다양한 목적을 위해 사용될 수 있다. 센서스 결과 보고에 있어서 지도는 대개 기술적(descriptive)인 목적으로 사용되지만, 지리학자나 인구학자는 다른 변수들 사이의 관계를 탐색하기 위해 지도를 사용할 수 있다.
⑥ 지도는 비교 분석을 고무시킨다. 기술적이든 설명적이든, 주제도의 주요 목적은 지리적 공간에 걸쳐 나타나는 현상들을 비교하는 것이다. '인구 밀도가 높은 지역은 어디인가'와 같이 다른 지역들을 비교할 수 있으며, '한 지역의 유아사망률이 다른 지역보다 높은가?'와 같이 지역 간 비교를 할 수 있으며, '한 지역에서 남녀 고용률은 어떻게 다른가?'와 같이 동일 지역의 다른 변수들 사이의 관계를 비교할 수 있다. 또한 '지난 센서스와 비교했을 때, 출생률의 증감은 어떠한가?'와 같이 시계열 비교를 할 수도 있다.

상업용 GIS와 데스크탑 컴퓨터의 폭넓은 보급은, 과거 지도학자들의 전유물이었던 주제도를 비전문가의 작업 영역으로 확장시켰다. 하지만 손쉽게 사용 가능한 맵핑 패키지들은 지도학적 원리에 어긋나는 지도들을 양산하여, 통계 지도의 양적인 증대를 가져온 반면, 질 높은 지도학적 산출물의 생산에 있어서의 그 유용성에 대해서는 의문이 제기되고 있다. 방위표나 축척과 같은 지도 레이아웃의 필수적인 요소들의 결여에서부터

적절한 주제도 표현 및 분류 방법에 대한 이해의 부족, 부적절한 해석 등 비전문가에 의한 주제도 제작에 따른 여러 문제들이 가시화되었다. 따라서 센서스 결과의 생산과 배포는 지도학자들이 담당하거나, 일반 비전문 담당자들로 하여금 일정기간의 지도학적 교육이나 연수를 통해, 양질의 디지털 통계 지도를 생산할 수 있도록 하는 것이 바람직하다 할 것이다.

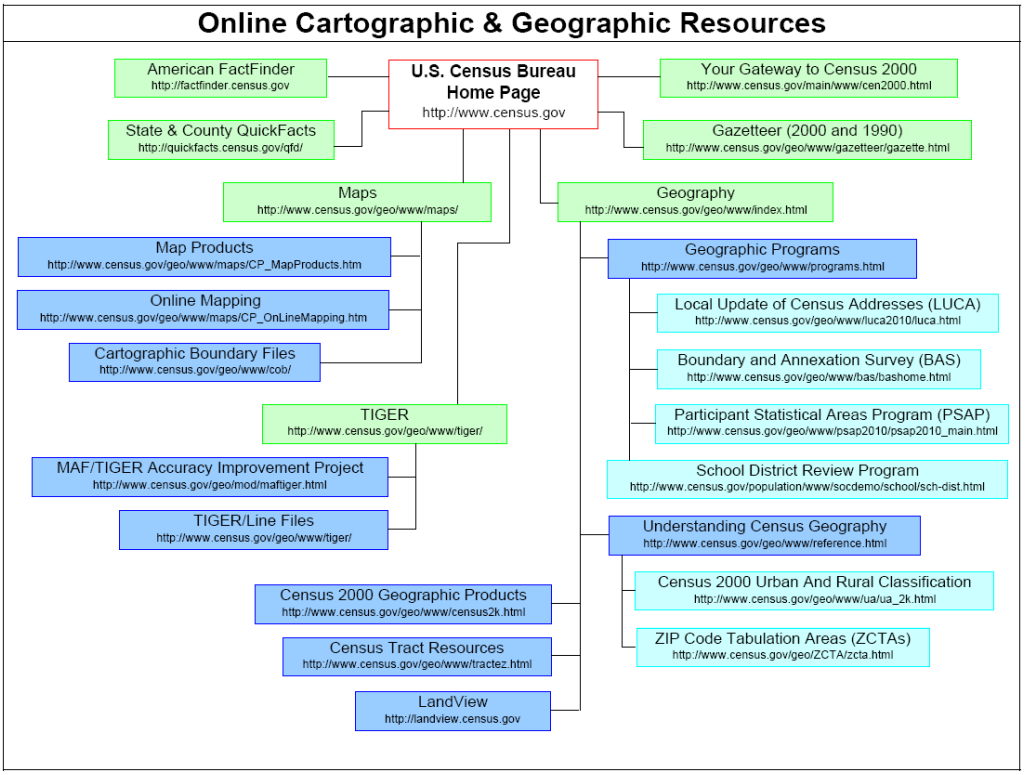
3) 다양한 형태의 지도 출판 및 유포

GIS를 통해 디지털 통계 지도뿐 아니라, 목적에 따라 다양한 형태의 지도학적 산출물이 출판될 수 있다. 표준 통계 단위에 따른 센서스 결과의 표준 출력 지도들에서부터, 센서스의 결과 또는 방법에 대한 기술을 목적으로 하는 보고서의 일부분으로 텍스트를 보완해주는 보조 자료로서의 지도, 또는 종이로 출력되어, 포괄적인 통계 지도를 포함하는 센서스 아틀라스 형태, 출력에 대한 비용을 절감하고, 사용자 지정 주제도 작성이 가능한 인터페이스를 제공하는 디지털 아틀라스 형태 등 여러 형태의 산출물들을 생산할 수 있다. 최근 특히 주목받고 있는 형태는 인터넷을 통한 주제도 서비스이다. 인터넷은 가장 빠르게 가장 많은 사람들이 센서스 데이터와 결과에 접근할 수 있도록 하는 수단이다. 디지털 이미지나 전자 문서와 같은 고정된 지도를 다운로드 할 수 있게 할 뿐 아니라 주제도 제작의 전반적인 프로세스에 사용자 권한을 부여하여, 사용자와의 상호작용을 극대화시킬 수 있는 동적인 맵핑 인터페이스를 제공할 수도 있다. 이러한 장점으로 인해, 현재 많은 국가 통계 기관들이 센서스 데이터를 인터넷 GIS를 통해 서비스하고 있다.

2. 주요 국가의 센서스 GIS

가. 미국

현재 미국 센서스에서 GIS가 통계와 얼마만큼 밀접하게 연관되어 있는지는 미국 센서스국 홈페이지를 통하여 제공되는 서비스 중에서 지리 혹은 지도와 관련된 서비스 이름 및 URL을 보여주는 [그림 6-19]를 통해 알 수 있다. 이들 중 일부는 센서스 조사 과정에 필요한 정보를 담고 있고 일부는 센서스 조사 결과의 서비스와 관련되며, 타이거처럼 일부 서비스는 두 범주 모두와 관련이 된다.



출처: www.census.gov/geo/www/tiger/webchart.pdf.

[그림 6-19] 미국 센서스 데이터의 지도학적·지리적 자원

미국 센서스국에서는 다양한 GIS 관련 프로그램을 운영하고 있는데, 이를 정리한 결과는 다음과 같다(<표 6-4>). 통계 속성 데이터뿐만 아니라 경계와 가로망에 관한 GIS 데이터가 함께 제공되고 있으며, American FactFinder 등을 통해 온라인상에서의 지도화(On-Line Mapping) 서비스가 제공되고 있다([그림 6-20]). 다양한 주제에 대한 스페셜 보고서들과 기 제작된 주제도 서비스를 통해 다양한 주제도를 서비스하고 있으며, 통계구 및 응용 통계구에 대한 자세한 참고 자료를 제공하고 있다.

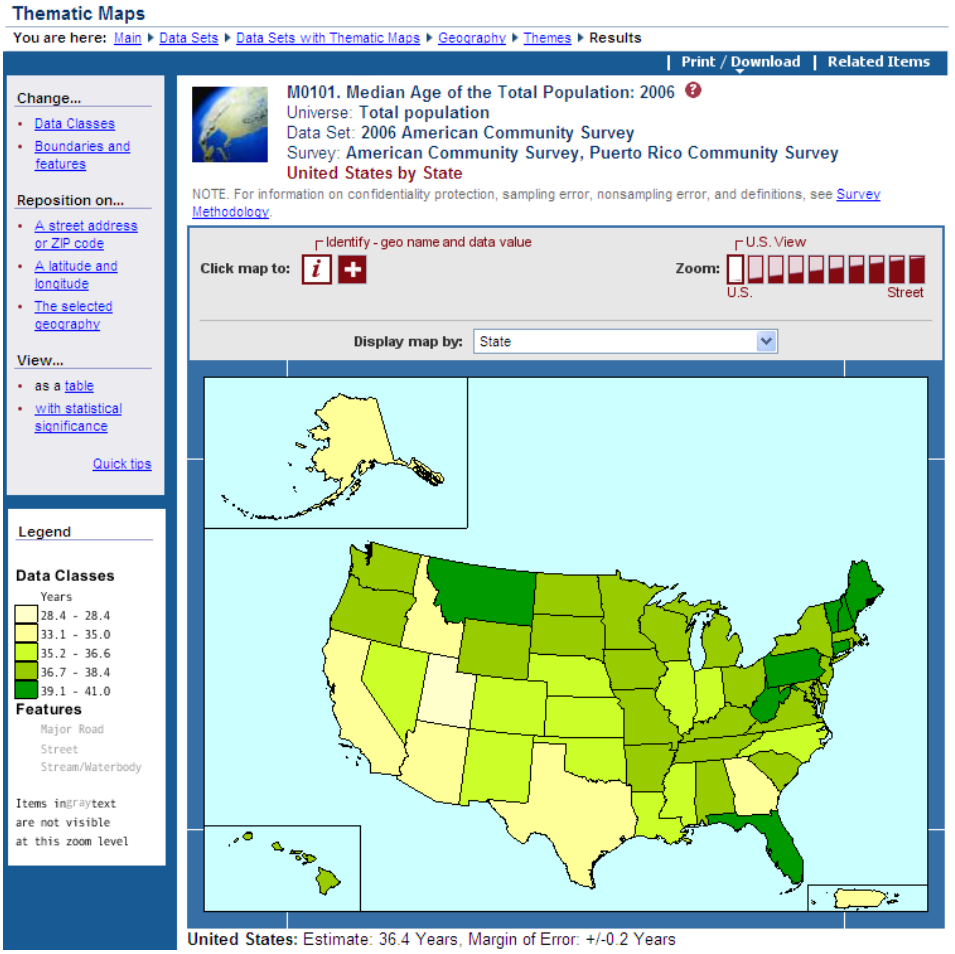
American FactFinder의 최소 제공 공간 단위는 카운티이며, 경계지도의 최소 제공 공간 단위는 센서스 블록이다. 블록 단위의 경계지도를 제공할 계획은 없음과 제공되는 경계지도들은 센서스국의 타이거 데이터베이스로부터 추출된 것으로, 이미지가 아닌 GIS에서의 사용을 위해 만들어진 것이라는 점을 명기하고 있다. 제공되는 경계 파일의 형태는 다음과 같다.

○ ARC/INFO EXPORT(.e00) format
○ ArcView Shapefile(.shp) format
○ ARC/INFO Ungenerate(ASCII) format

〈표 6-4〉 미국 센서스국의 GIS 관련 프로그램

항목	세부 프로그램	내용
Geographic Program	2010 Census Local Update of Census Addresses(LUCA) Boundary and Annexation Survey(BAS) School District Review Program 2010 Census Participant Statistical Areas Program(PSAP	프로젝트형 프로그램 운영
타이거 관련 Program	MAF/TIGER Enhancements Program Public TIGER/Line Shapefiles TIGER/Line Files TIGER-Related Products	타이거 관련 GIS 파일의 관리와 제공
LandView 6	A Viewer for the Environmental Protection Agency, U.S. Census Bureau, and U.S. Geological Survey Data and Maps	GIS Viewer 프로그램의 제작 및 배포
Maps and Mapping Resources	완성된 지도의 이미지 Boundary Files On-Line Mapping	GIS 경계 파일의 제공 및 지도화툴 제공
Census Tract Resources	Census Tract Number for an Address Census Tract Outline Maps Locating Census Tracts	센서스 트랙 파일 및 연계 툴 제공
Census Geography Reference Resources	Hierarchical Diagrams Geographic Glossaries and Guides Geographic Area Criteria Census 2000 Reference Files	센서스 집계구 및 공표구에 대한 참고 사항 제공
Census 2000 Geographic Products and Information	Congressional Districts Census 2000 Maps Census 2000 Data Files Census 2000 Reference Resources	센서스 2000과 관련된 결과와 참고 정보 제공

출처: http://www.census.gov/geo/www/index.html.



출처: 미국 센서스국 홈페이지(http://www.census.gov)

[그림 6-20] 미국의 인터넷 GIS 서비스인 American FactFinder

나. 영국

1) 1991년 센서스와 GIS

영국 통계청(Office for National Statistics: ONS)[17]은 10년마다 센서스를 실시하는데, 1991년의 센서스를 위해, 측지국(Ordnance Survey)이 제작한 대축척의 종이 지도를 바탕으로 센서스 담당자들이 직접 손으로 조사구 지도를 제작하였다. 조사구는 일차적으로

17) 1996년에 기존의 Office of Population Censuses와 Central Statistical Office를 통합하여 영국 통계청이 만들어졌다.

조사원들의 효율적인 자료 수집과 분배를 촉진하기 위한 지리적 범위이기 때문에, 조사원들 사이에 거의 같은 작업량이 분배되도록 해야 하며, 물리적 접근성을 높일 수 있어야 한다. 따라서 형태와 크기가 데이터 수집을 위한 요구에 따라 결정되는데, 또한 국가적 일관성을 위해 Ward, Parish/Community와 같은 상위 공간 단위들에 포함되어야 한다. 이러한 조건들을 만족시키는 형태로 조사구가 디자인되었으며, 1991년에는 England and Wales에서 116,919 기본 조사구와 4,840 특별 조사구(지리적 경계 없음)가 설정되었다(Martin, 2002). 이렇게 수작업으로 제작된 조사구는 인구나 주택의 지역적 변화에 따라 다시 제작되어야 하는 비효율성을 가지고 있었는데(Martin, 1997)[18], 1980년대 이후 급격하게 성장한 GIS를 이용하여 디지털화할 수 있었다. 하지만 1991년 센서스까지도 여전히 실제 조사구 설정 과정은 수작업으로 수행되었다.

2) 2001년 센서스와 GIS

1991년 센서스의 조사구는 센서스의 조사와 공표에 있어 단일 단위로 사용되었다. 이러한 단일 단위의 사용은 인구 규모의 변화에 따른 조사구 경계의 변화, 조사구간 인구 편차의 존재, 개인정보 보호가 어려운 인구 규모의 조사구 존재, 사회적 동질성이나 규칙적인 형상 조건의 불합치(변필성 외, 2006), 당시 보편적으로 이용되었던 우편 구역과의 불일치, 다양한 크기의 공표구역의 설정의 어려움 등과 같은 여러 문제를 낳았다. 하지만 1991년 센서스의 경우, 조사구의 획정이 수작업으로 이루어졌으므로, 많은 문제점이 있었음에도 불구하고 조사구와 별도로 공표구역을 획정·운영하는 것이 물리적으로 불가능했다.

하지만 1990년대 이후 GIS의 발전과 컴퓨터 기반 속성 자료 입력과 경계선 입력·획정 기법의 발전으로 인해 특수한 목적에 맞게 센서스 공표구역을 획정할 수 있게 되었다(Openshaw and Rao 1995; Martin, 1998). 즉 GIS의 발전을 바탕으로 개별 주택의 주소에 기반을 둔 데이터(address referenced data)를 활용할 수 있게 됨으로써, Output Area(OA)라는 소지역 센서스 공표구역을 효과적으로 획정할 수 있게 되었다. GIS는 2001년 센서스의 조사구과 공표구역의 디자인 모두에 중요한 역할을 하였다.

2001년 센서스의 두드러지는 특징은 기존의 단일화된 조사구과 공표구역을 분리하여 새로운 공표구역인 OA를 생성했다는 것이다. 조사구 설정은 구역 경계를 디지털화하는 것을 넘어 전체적인 디지털 환경에서 작업이 이루어졌다. 측지국의 ADDRESS-POINT[19], 경계 라인, 래스터 지도들을 자료원으로 상업용 GIS 소프트웨어 중 하나인

18) 77,000개의 기본도와 약 400,000개의 복사본에 새로운 경계를 직접 수작업으로 표시해야 했다(Martin, 1997).

19) 이는 영국 체신공사(Royal Mail)의 우편코드 주소 파일(Postcode Address File: PAF)에 기반한 디지털 주소 파일이다.

ESRI社의 ArcInfo에 기반하여 GAPS(Geographical Area Planning System)를 개발하였다. 잉글랜드와 웨일즈에서는 평균 200세대, 450명 규모의, 116,895개의 새로운 조사구 데이터베이스를 구축하였다. 1991년에 디지털화된 조사구를 기반으로, 각 조사구 내에 모든 주소의 포인트(address point)를 할당했으며, 현 조사구 내의 주소 포인트수를 재계산하여 적절한 조사 업무량에 따라 재조정이 필요한 조사구를 확인하고, 경계를 수정하였다. 이렇게 완성된 조사구는 여러 장의 종이 지도를 사용할 필요 없이, 기존의 디지털 경계, 래스터 지도 등을 하나의 스크린에 중첩하여, 하나의 지도로 제작될 수 있었다.

따라서 2001년에 조사구는 조사 목적만을 위해 사용되었고, 공표구역은 우편구역 기반의 OA로 대체되었다. OA는 스코틀랜드에서 1981년 센서스 공표단위로 처음 소개되었고, 이후 잉글랜드와 웨일즈 그리고 북아일랜드에서는 2001년 센서스에 처음 사용되었다. OA는 몇 개의 인접한 우편코드 단위(일반적으로 15개 주소 포함)를 묶어서 조사구보다 작은 크기로 설정되었기 때문에 소지역 분석에 적절히 이용될 수 있었다. OA는 유사한 인구 규모, 사회적 동질성, 구역 형태, 사생활 보호를 위한 최소 인구 규모 등의 기준에 의해 설정되었다. 최소 OA 크기는 40세대, 100명 규모(Scotland는 20세대, 50명 규모)로 설정되었고, Ward 또는 Parish의 통계 단위에 포함되도록 설계되었다. 잉글랜드에 165,665개, 웨일즈에 9,769개, 북아일랜드에 5,022개, Scotland에 42,604개의 OA가 존재한다. OA의 설정은 자동화된 컴퓨터 알고리즘에 의해 생성되었다.[20)]

3) 인터넷 GIS 서비스

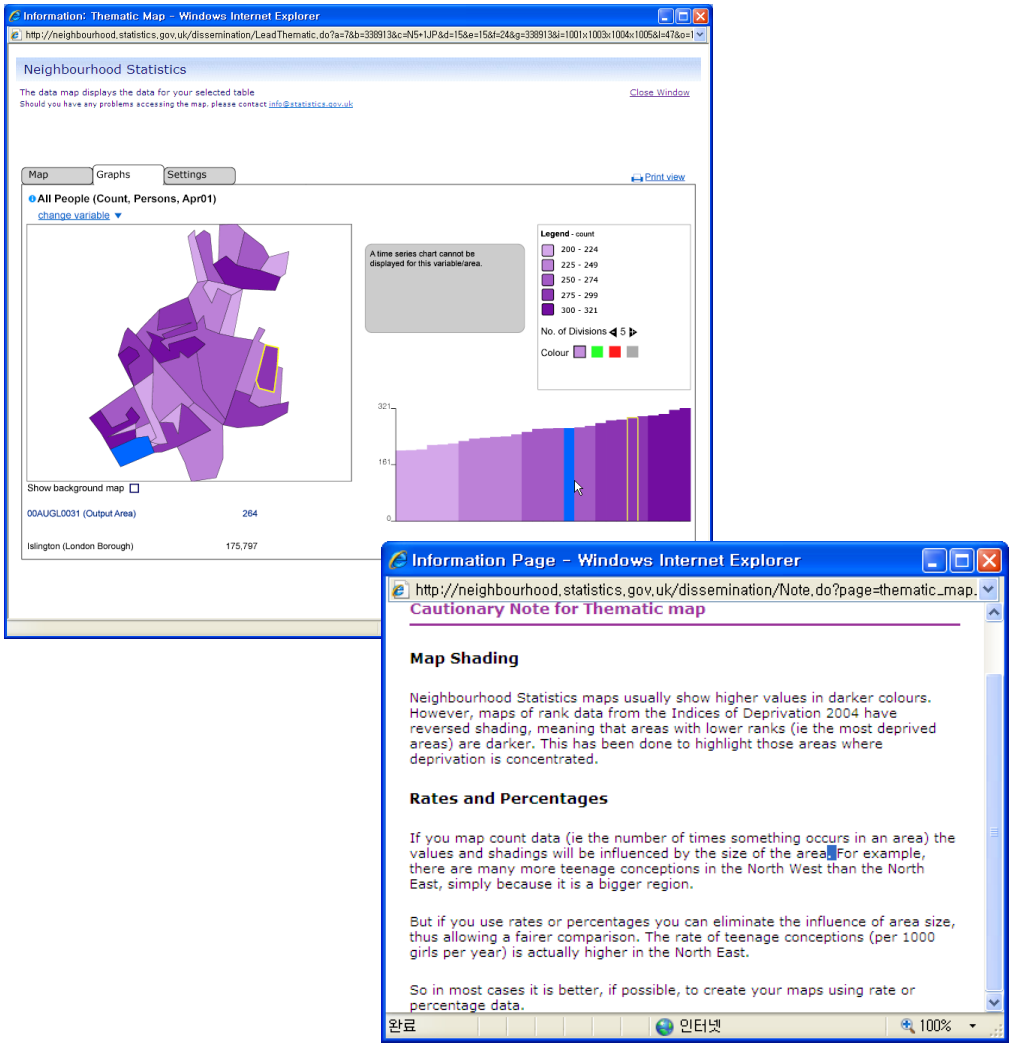
영국 통계청의 Neighbourhood Statistics(NeSS)는 웹 기반의 통계지리시스템으로 여러 데이터에 대한 주제도를 생성해주며, Ward/Parish 등의 상위 행정 단위뿐 아니라 표준 공표구역인 OA 및 Super OA[21)]에 이르는 공간통계를 서비스하고 있다. NeSS는 소지역 통계에 대한 이용 촉진을 위해 2001년부터 5년간의 개발 과정을 통해 시행되었으며, 2006년에 성공적인 평가를 받아 영국 통계청의 서비스 영역으로 귀속되었다. 행정경계의 변화에 대한 영향을 최소화하는 보다 안정적인 소지역 통계 생산을 위해, OA, Super OA 등으로 구성된다. 2001년 센서스를 포함하여, 주요 시설물에 대한 접근성, 교통사고, 화재, 범죄 등 소지역의 다양한 사회경제학적 토픽들에 대한 주제도 및 다운로드 서비스가 제공되고 있다. NeSS는 다양한 수준의 주제도 제공에 있어, 색상, 계급의 수 등을 사용자가 임의로 변경할 수 있는 인터페이스를 제공하여, 주제도 제작의 유연성을 부여하고 있다.

---

20) Openshaw(1977)에 의해 개발된 Automated Zoning Procedure(AZP)를 통해 생성되었으며, 우리나라의 자동화된 집계구 설정(건교부, 2007) 역시 이 알고리즘과 영국 OA에 사용된 기준을 참고로 디자인되었다.

21) 표준 OA와 local authorities의 중간 크기로 설정되어, Neighborhood Statistics의 디폴트 공간 단위로 사용되고 있다.

특히 주목할 만한 NeSS의 디자인은 팝업 형태의 노트를 통해 지도학적 표현에 대한 주의 사항을 설명하고 있다는 점이다. 잘못된 지도학적 표현으로 데이터의 왜곡이 발생하는 것을 사전에 숙지시켜 보다 정확한 주제도의 제작을 고취시키고 있다([그림 6-21]).



주: 좌측의 지도뷰와 우측의 하단의 분포 그래프는 동적으로 연결되어 있으며, 우측 상단의 데이터 분류 프레임웍을 통해 즉각적인 주제도의 변경이 가능하다. 주제도의 표현, 특히 면적에 영향을 받는 절대수(예, 지역의 총인구수)의 단계구분도(choropleth map) 표현 시 주의 사항을 설명해주고 있다.

[그림 6-21] NeSS의 주제도 서비스와 지도학적 주의 사항을 표시하는 팝업 노트

영국의 웹기반 통계지리시스템은 통계의 주제도 표현에 있어 지도학적 요소나 개념들에 대한 이해에 기반하고 있으며, 간편한 마법사 형태의 인터페이스와 주제도 제작 시의 사용자 지정 옵션을 확대함으로써, 사용자와의 상호작용 측면을 강조하고 있다.

다. 한국

국내에서는 행정 단위가 일반적인 통계 자료의 공표 단위로 활용되었기 때문에, 소지역 통계 혹은 통계공표 단위에 관한 논의와 관련 연구가 활발히 이루어지지지 않았다(강영옥 외, 2007). 대부분 통계의 최소 행정 단위인 읍면동은 여러 가지 문제점이 있다. 우선 행정경계가 바뀌고 GIS적인 이력관리가 안되는 문제점이 있고, 인구 및 산업구조의 변화에 따라 낙후된 지역과 광역화되어가는 지역의 인구가 3,000명 미만에서 30,000명 이상으로 규모의 격차가 커지고 있으며, 공간 단위가 너무 커서 동질적인 지역으로 보기 어려운 실정이다. 읍면동 단위의 공표 통계가 다원화되어가는 지역의 특성을 대표적으로 반영할 수 있는 현실 지표로서, 점차 전문화되어 가는 수요자의 욕구를 충족시킬 수 있을 것인가에 대한 비판적 고려가 요구되는 시점이다(강영옥 외, 2008). 따라서 시간이 지나도 자료를 이력관리 할 수 있는 표준화된 소지역통계 단위가 필요한데, 이러한 표준화된 통계 단위의 획정은 통계청에서 국가의 근간이 되는 통계를 제공하고 있기 때문에 통계청 업무와 긴밀하게 연관된다고 볼 수 있다. 통계청에서는 행정동 단위가 지나치게 크다는 지적을 인식하고, 관련 연구를 지속적으로 수행한 바 있으며(강영옥 외, 2008), 공간 데이터베이스 구축과 서비스에 있어 지속적인 개선 작업을 수행하고 있다.

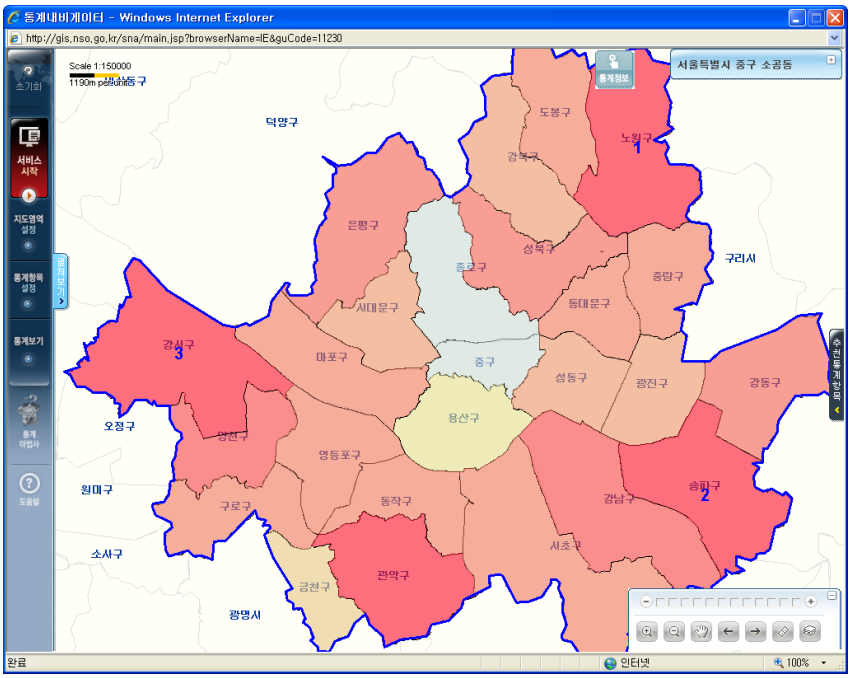
거처 포인트 단위로 기본 GIS 데이터가 구축되어 있으며, 이를 기반으로 기초단위구, 조사구, 집계구 등의 통계 단위를 생산하고 있다. 2008년 기준으로 거처 포인트 13,459,433건 기초단위구 331,596건의 자료가 구축되어 있다(통계청, 2008b). 기초단위구는 도로, 하천, 철도, 산능선 등과 같은 준항구적인 명확한 지형지물을 이용하여 지도상에 구획한 최소 단위구역으로 2000년 인구주택총조사부터 사용되었는데, 이는 조사구 획정의 기본 단위가 된다.[22] 자료 생성과 업데이트를 위한 기초 자료로 다양한 GIS 데이터가 이용되고 있다(<표 6-5>). 2009년 5월 이후에는 전국을 대상으로 집계구에 대한 인터넷 GIS를 서비스 중에 있다([그림 6-22]).

22) 우리나라의 조사구는 통계조사 시 그 대상이 되는 지역의 조사담당을 명확히 하고 조사원의 근무량을 적절하게 배분하는 것을 목적으로, 일정 규모의 가구수를 기준으로 분할한 구역으로 보통 하나의 조사구는 60가구 정도의 규모로 구성된다(통계청, 2008a).

〈표 6-5〉 통계청에서 활용 중인 지리정보 데이터 현황

기간	구분	내용
2000년 수치지도 도입	인구주택 사업체	• 국립지리원(현 국토지리정보원) 및 지자체에서 제작한 수치지도 입수 • 2000년 센서스 조사구 설정용 기본도 작성 • 현장 확인을 통한 보완 및 조사구 설정
2001~2003년	-	• 현지 확인을 통한 지도의 수정, 보완 • 국토지리정보원에서 새로 구축한 지역의 경우 • 도엽별 지도를 입수하여 적용
	-	• 센서스 2000기준 추가 자료 생성 • 거처포인트 생성 • 기초단위구 경계 생성 및 보완(2001년, 2002년, 2003년)
2004년	인구주택	• 일부 지자체에서 구축된 새주소 지도의 입수 • 새주소 지도에 맞추어 지형지물을 제외한 자료 수정 • 새주소 미획득 지역은 수치지도 수정, 보완
2006년~현재	인구주택	• 새주소 지도와, 수치지도를 수정·보완 사용 중
	국가표준 수치활용지도 제작을 위한 계획 수립	• LMIS(Land Management Information System), 새주소 자료를 입수 • 2005년 인구 및 사업체 센서스 자료를 이용해 자료 수정 및 보완

출처: 통계청, 2008b.



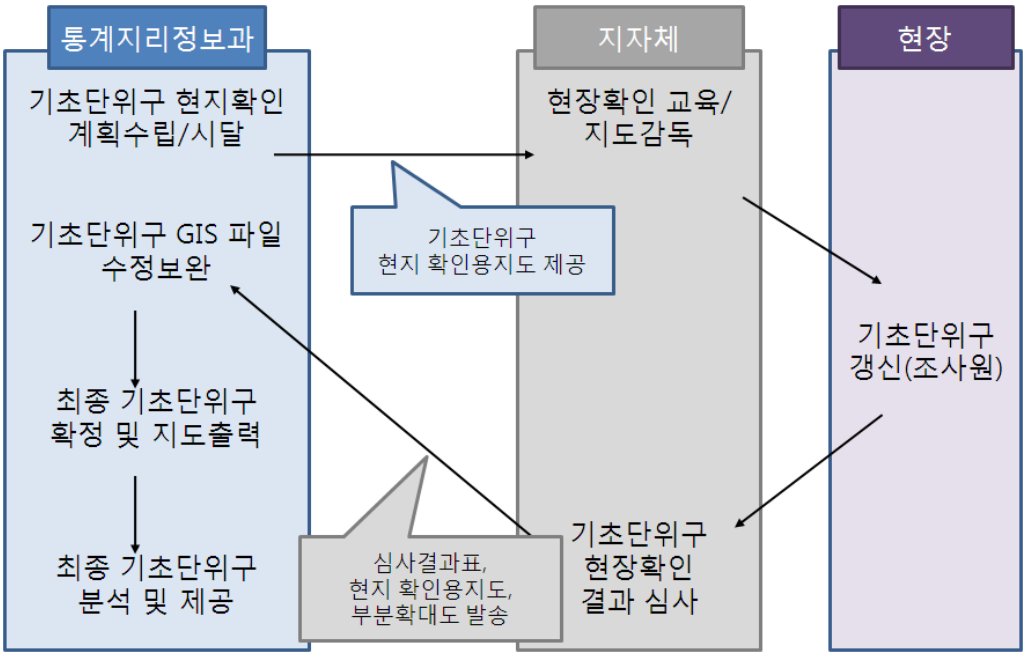
출처: 통계청 홈페이지(http://gis.nso.go.kr/).

[그림 6-22] 우리나라의 인터넷 GIS 서비스

기초단위구의 구획 방식은 다음과 같다(강영옥 외, 2007).

- ○ 가구수에 관계없이 읍면동 내에서 먼저 대도로, 대하천, 산맥 등과 같은 주요 지형지물을 이용하여 대구역을 구획하고, 그 다음 명확한 지형지물을 이용하여 소구역을 구획한다.
- ○ 시가지 조성으로 인하여 블록화된 지역은 가구수에 관계없이 블록 단위로 구획하고 아파트 지역은 건물 동 단위로 구획한다.
- ○ 시가지 주변지역이나 농촌지역은 가급적 자연부락 단위로 구획하는데, 명확한 지형지물이 없는 경우에는 통리 경계를 이용하여 구획한다.
- ○ 시가지 변두리지역 또는 농촌, 산간지역은 면적이 다소 크더라도 가급적 자연부락이 분할되지 않도록 구획하며, 명확한 지형지물 이용이 곤란한 경우 통리 경계를 활용한다.
- ○ 도서 지역은 섬 단위로 구획하되, 자연부락이 2개 이상인 경우 분할해서 기초단위구를 획정한다.
- ○ 개발이 이루어지고 있는 지역은 전체를 하나의 기초단위구로 설정하며 개발이 완료된 후에 구획 기준에 의해 분할한다.

한편 통계청의 기초단위구 관리 업무 체계를 관리 주체별로 구분하여 도식화한 결과는 다음과 같다([그림 6-23]).



[그림 6-23] 기초단위구 관리업무 체계

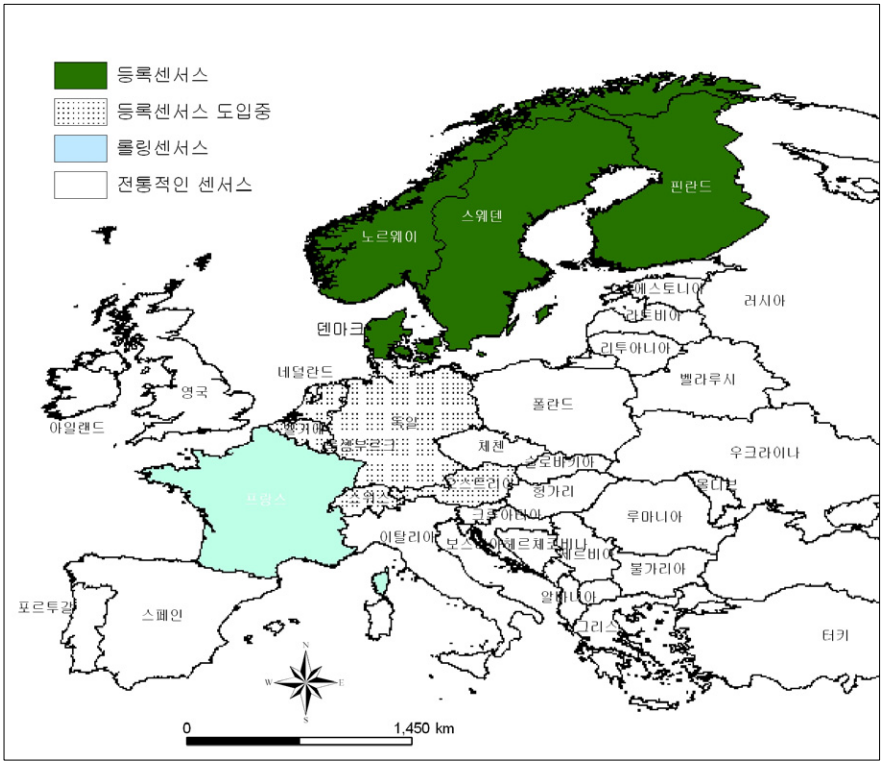
통계 공표의 기본 단위인 집계구의 설정 원칙은 다음과 같다(강영옥 외, 2007).

- ○ 집계구는 사용자에게 유용해야 한다.
- ○ 행정구역 단위 등 다른 분석체계와 통합이 쉬워야 하며, 이를 위해서는 통계 단위의 계층성이 확보되어야 한다.
- ○ 집계구의 형태가 안정적이어야 하고, 집계구 변경 시에는 시계열 분석이 가능하도록 집계구의 이력을 관리해야 한다.
- ○ 집계구는 사회경제적인 동질성을 확보해야 한다.
- ○ 범국가적 일관성을 유지해야 한다.
- ○ 통계 자료 공표시 개인의 사생활 보호를 고려해야 한다.
- ○ 집계구는 누구나 쉽게 이해할 수 있도록 단순하여야 한다. 집계구의 경계는 누구나 쉽게 식별이 가능하도록 현저한 지형특성을 기준으로 구분하여야 한다.

제4절 등록부 기반 센서스 실시 국가에서의 GIS

1. 등록부 기반 센서스와 GIS

최근 일부 국가에서 센서스 조사 방법의 다양화가 이루어지고 있는데, 이는 조사비용의 확보 문제, 응답부담과 응답자의 반발에 따른 응답률의 감소 등 조사환경의 변화와 주로 관련된다(조순기 · 최은영, 2009). 등록부에 기반한 센서스는 행정 등록부를 통하여 데이터가 수집되고, 관리되고 조작되는 시스템이다. 등록제도가 발달한 일부 북유럽 국가를 중심으로 등록센서스로의 전환이 이루어졌거나 전환이 검토되고 있다. 덴마크 · 핀란드 · 노르웨이 · 스웨덴 4개국에서 완전한 형태의 혹은 불안전한 형태의 등록센서스가 실시되고 있다. 이 4개 국가는 주로 인구 규모가 작고 사회보장제도가 잘 되어 있다는



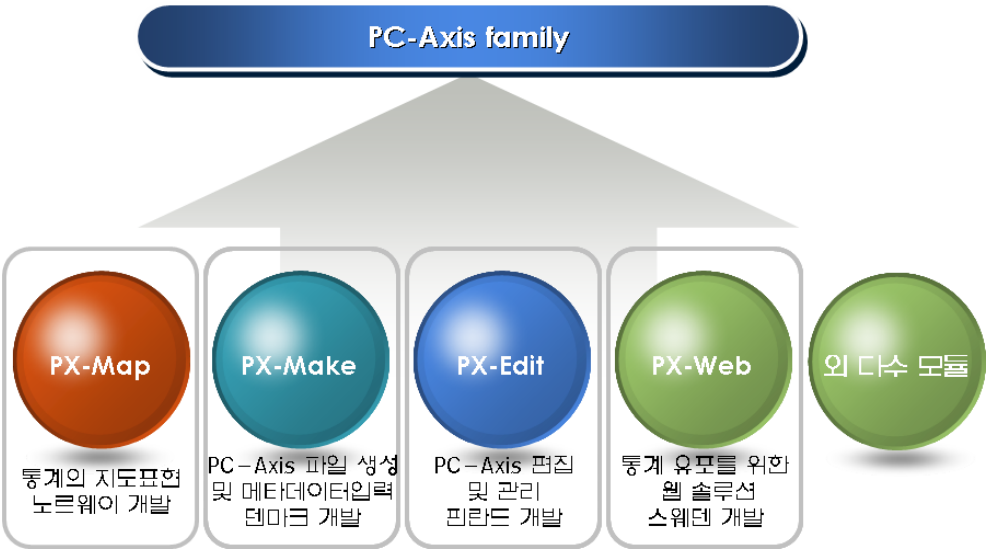
자료: http://www.statistik-portal.de/Statistik-Portal/en/Zensus/en_zensenA.asp.
출처: 조순기 · 최은영, 2009.

[그림 6-24] 유럽의 국가별 센서스 실시 현황

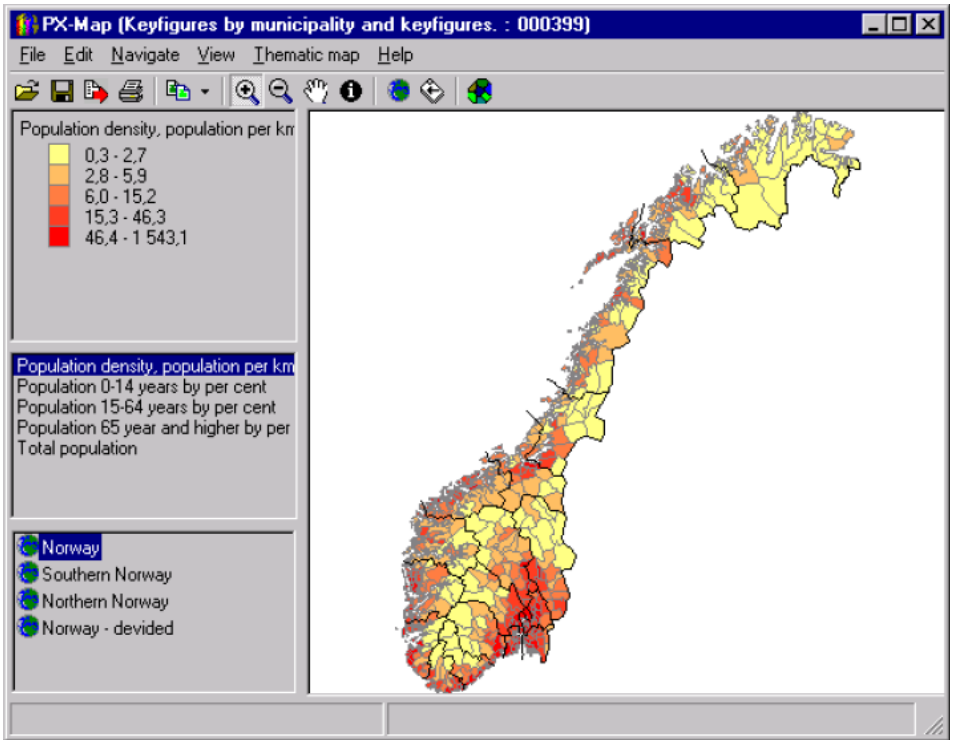
특징을 가진다. 이들 국가와 인접해 있는 스위스・네덜란드・룩셈부르크・독일・벨기에・오스트리아 6개국에서는 등록센서스를 도입 중이다. 2007년을 기준으로 유럽 49개 국가의 센서스 조사 방법을 살펴본 결과는 다음과 같다([그림 6-24]).

등록부에 기반한 데이터베이스 간의 통합은 개별 자료의 절대좌표 값을 제공할 수 있는 지리정보시스템의 지원으로 보다 효과적이고 강력해질 수 있는데(김형석 외, 2003), 모든 북유럽 국가에서는 지리정보시스템의 발전과 함께 개별 건물의 포인트에 기반한 통계 시스템을 발전시켜왔다(UNECE, 2007). 등록부에 기반한 센서스를 실시하고 있는 국가들의 통계청에서는 건물의 좌표를 수집하고 이들 자료를 등록부에 입력・업데이트하는 작업을 수행해 왔다. GIS의 연계 시스템을 사용해서 개별 통계 단위의 정확한 위치를 할당할 수 있도록 모든 건물의 위치가 정확한 x, y 지도 좌표를 사용해서 특정화 되어 있다. 좌표를 가지고 있는 건물 기반의 코드 체계는 신뢰할 수 있고 유연한 통계 지역을 만들어 내기 위한 튼튼한 기반을 제공한다. 이와 같이 건물에 대한 좌표 체계의 도입이 보다 유연한 통계 지역의 설정을 가능하게 하는 이유는, 대부분의 GIS 패키지들에 폴리곤(통계 지역) 내의 포인트(건물)에 관한 폴리곤 내 포인트(point-in polygon) 분석 기능이 포함되어 있기 때문이다(조순기・최은영, 2009). 등록부를 기반으로 센서스를 실시하는 국가에서의 통계지리시스템은 통계의 속성 데이터와 공간 데이터를 결합하여 각종 통계 생산 및 서비스를 제공하고 있다는 점에서는 같지만, 구체적인 결합 및 구축 방식은 국가마다 조금씩 다르다.

등록부에 기반한 센서스를 실시하고 있는 북유럽의 4개국에서는 공통의 GIS 패키지인 PC-Axis를 공동으로 개발하여 사용하고 있다([그림 6-25]). 1990년 인구 센서스의 데이터를 표현하고, 유포하기 위해 스웨덴에서는 무료 프로그램 패키지인 PC-Axis를 초기 개발했는데, 국제 PC-Axis 레퍼런스 그룹이 다른 참여자들과 함께 이후의 개발을 이끌었다. 현재 북유럽 국가, 스페인 등 유럽의 여러 나라를 포함해서, 세계 30여 개 국가에서 이를 국가 통계의 공표를 위한 공통된 통계 데이터베이스 시스템으로 사용하고 있다. PC-Axis 프로그램은 통계의 공표 과정과 통계 데이터 생산 및 처리의 모든 과정에 적합한 많은 하위 모듈들을 포함한다. 이 패키지는 국제적인 공동 작업을 통해 만들어졌는데, PX-Map은 통계의 지도 표현을 위해 노르웨이에서 개발되었고([그림 6-26]), PC-Axis 파일을 생성하고, 메타데이터를 입력하는데 필요한 PX-Make는 덴마크에서 개발하였다. 핀란드에서는 PC-Axis를 편집하고 관리하기 위한 툴로 PX-Edit를 개발하였고, 스웨덴은 웹 솔루션인 PX-Web을 개발하였다(보다 자세한 정보는 웹페이지 http://www.pc-axis.scb.se/ 참조).



[그림 6-25] 북유럽 국가의 PC-Axis 소프트웨어와 하위 모듈



[그림 6-26] PX-Map의 인터페이스

2. 북유럽 국가의 센서스 GIS

가. 스웨덴

1) 통계지리시스템의 구축

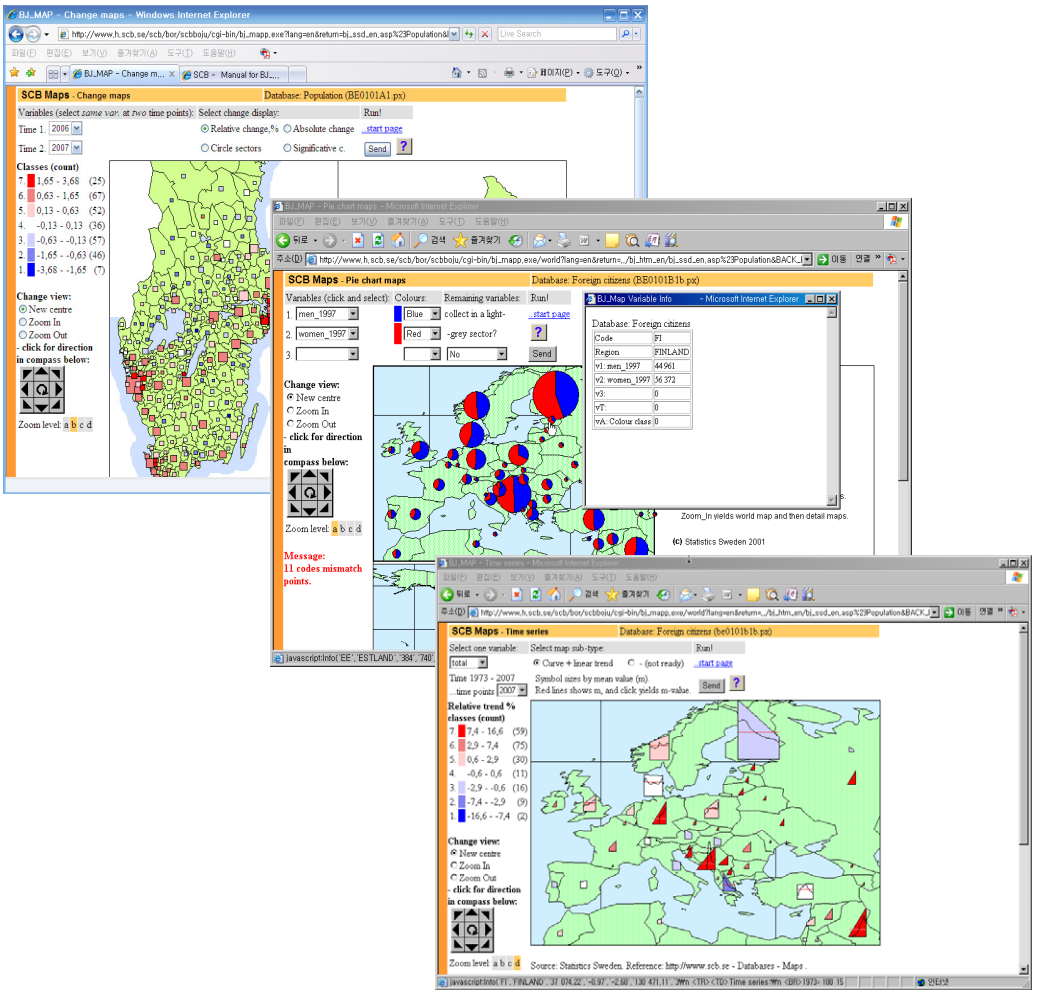
스웨덴의 통계지리시스템은 행정등록부에 기반하고 있으며, 포인트 기반으로 개발되었다. 마이크로 통계의 생산을 위해 7자리 좌표로 위치를 표시하고 있으며, 가장 작은 센서스 단위는 1미터 정확성을 가지는 포인트이다(Nael and Nordström, 1998). 스웨덴의 토지조사국은 1960년대에 이미 디지털 토지 정보 처리를 시작했고, 모든 개별 필지 단위들을 디지털화하였다. 여기에는 개별 필지들과 건물들의 x, y 좌표를 포함하고 있다.

스웨덴 통계청은 county, municipality의 좌표를 제공하고 있으며, 다른 통계 단위들, 예를 들어, 우편구역, 그리드(grid), SAMS(Small Area Market Statistics) 등의 디지털 경계는 유료로 제공하고 있다.[23] 이미지 파일은 웹에서 다운로드 가능하다. 부가적으로 여러 통계 단위별 통계, 시장 프로필(market profile), 거주지 프로필(residential profile)과 같은 GIS 부가가치 생산품을 유료로 제공하고 있다.

2) 인터넷 GIS 서비스

스웨덴 통계청의 인터넷 GIS 서비스인 BJ-Map은 주제별 속성 데이터와 그래프, 주제도를 municipality 단위로 제공하고 있다. 스웨덴 인터넷 GIS 서비스의 주요 특징은 서클지도(circle map), 파이차트지도, 단계구분도, 시계열지도(time series map)와 같은 다양한 형태의 주제도를 제공하고 있다는 것이다([그림 6-27]). 또한 각 지역에 대한 상세 속성정보를 마우스 클릭을 통해 살펴볼 수 있다. 또 하나의 주요 특징은 등간격, 고정비율과 같은 계급 분류 방법, 계급의 수, 심벌의 색상 등을 사용자가 변경할 수 있기 때문에 유연한 주제도 제작이 가능하다는 것이다. 하지만, 출력을 위한 레이아웃이 없고, 확대/축소, 이동과 같은 탐색 툴이 없어, 사용자 친화적인 인터페이스를 제공하지 못하는 단점이 있다. 또한 지도학적 측면에서 축척을 표시하지 않아 지리적 스케일을 파악하기가 용이하지 않다. 통계 및 공간 데이터의 원자료를 인터넷을 통해 직접 다운로드할 수 없는 한계점도 가지고 있다.

23) 등록부 기반 센서스 실시 국가에서의 그리드 데이터에 관해서는 다음 절에서 보다 상세히 살펴보도록 한다.



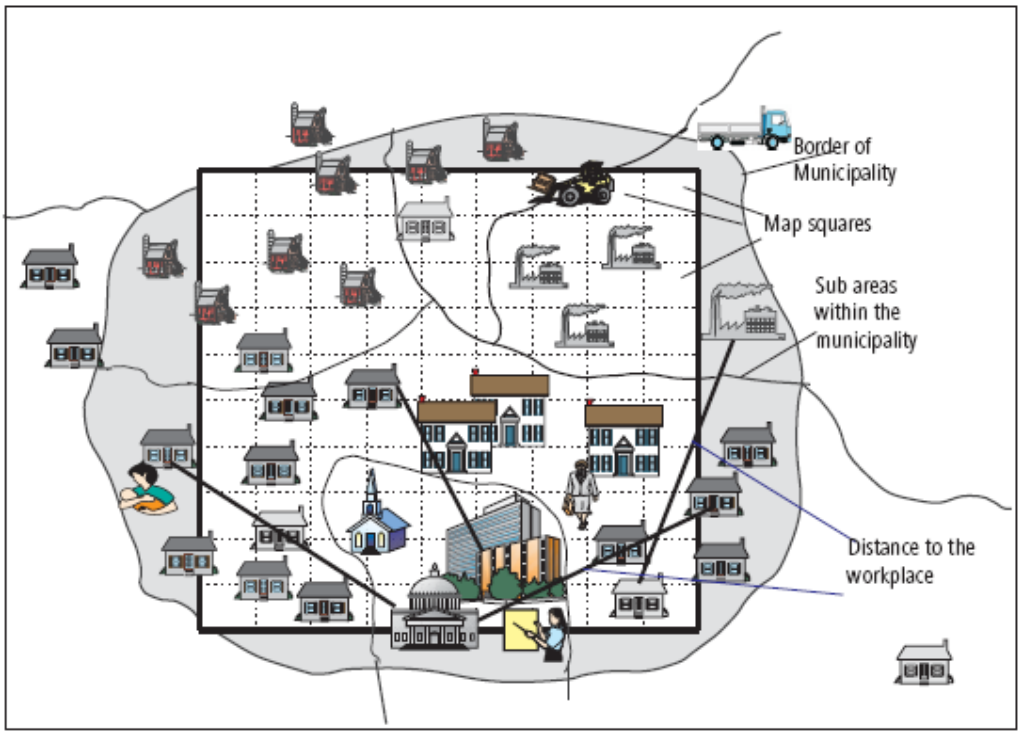
[그림 6-27] 스웨덴의 인터넷 GIS 서비스

나. 핀란드

1) 통계지리시스템의 구축

핀란드 통계청(Statistics Finland)의 GIS의 전통은 1970년 인구 센서스로부터 시작되었다. 이 때 건물의 좌표 데이터가 시당국으로부터 처음으로 수집되어 데이터베이스화 되었다. 통계지리시스템 역시 건물의 위치 좌표를 이용한 포인트 기반으로 구축되었다. 다음 [그림 6-28]은 건물의 개별 좌표 단위로부터 지역 통계를 위한 지리 단위들과의 연결을 보여주고 있다. 초기의 주요 서비스는 디지털 지도 데이터, 주제도 출력, 그리드 단위

의 공간 통계 데이터, 공간 분석 결과, 지오코딩, 컨설팅 및 교육, GIS 프로그램의 재판매, 핀란드 통계 연보의 제공 등이었다(Laurila, 2001).



출처: Statistics Finland, 2004.

[그림 6-28] 건물의 입력 좌표와 다른 공간 단위와의 연계

2000년 센서스를 위하여 핀란드 통계청은 도시 기능지역인 '로컬리티(locality)'를 GIS를 이용하여 분류하였다.[24] 핀란드 통계청은 주제별 속성 데이터를 온라인으로 무료로 제공하고 있지만, 그 외 다음과 같은 유료 상품들도 생산하고 있다.

- ☐ AreaOnline: 유료 지역(region) 정보 서비스, 주제도와 그래프를 포함한 이미지 제공(19,000 이미지)
- ☐ Municipal facts: 연간 2회 업데이트, 소프트웨어 패키지, municipality 단위별, 200가지 속성 정보 제공, 시별 데이터 보기 가능, 두 개의 시 비교 가능, 시별 테이블

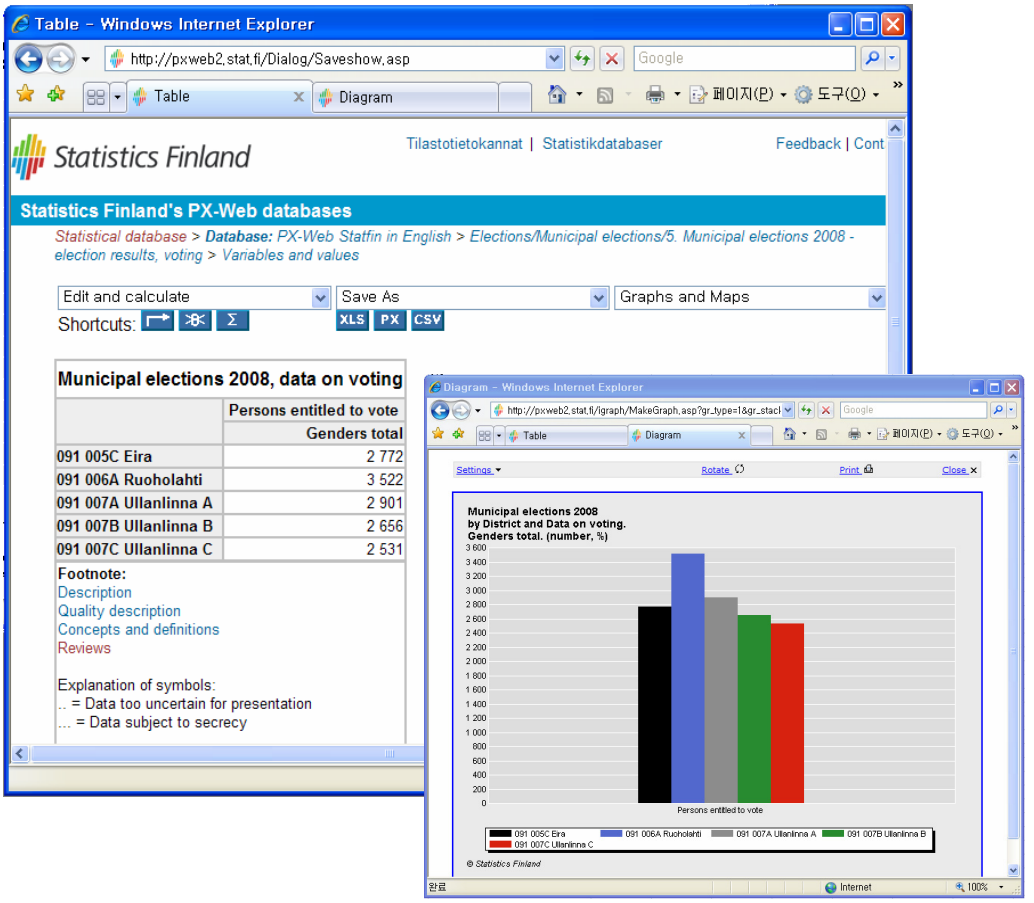
24) 핀란드는 인구밀도가 낮아서 유럽적인 의미에서 '도시'로 분류되는 곳은 많지 않고, 대신 로컬리티가 인구 중심의 의미를 가지고 있다.

생성 및 그래픽 제공

☐ SuomiCD: 우편 코드 지역구(2741)와 municipality(431)에 따른 국지적 데이터 제공, 지역의 거주자(inhabitants)와 가구의 소득・직업・직장과 거주지 사이의 거리・자동차 등에 관한 정보 제공

☐ Grid database: 지도의 그리드에 의해서 계산된 지리 참조된 통계 데이터베이스로 250m x 250m 와 1km x 1km 의 두 가지 해상도의 제공, 인구 구조・교육・생산활동・소득・건물・직장 등의 변수 선택 가능, 연간 업데이트

☐ Regional net service: 하위 지역 단위들과 지방정부에서 생산한 지역 통계에 대한 수요를 만족시키기 위해 개발된 유료 인터넷 서비스로 지방정부 홈페이지에 직접 링크됨



[그림 6-29] 핀란드의 인터넷 GIS 서비스

2) 인터넷 GIS 서비스

핀란드의 인터넷 GIS 서비스는 StatFin으로 불리우며, 북유럽 국가의 공통 통계 소프트웨어인 PC-Axis 데이터베이스의 하위 기술인 PX-Web을 사용하여 개발되었다([그림 6-29]). 통계 데이터를 엑셀, 텍스트 파일, PC-Axis 전용 파일, dbase 파일 등의 여러 형태로 무료로 다운 받을 수 있다. 지도는 웹기반 통계 데이터베이스와 직접 연동되어, 그래프와 더불어 통계의 표현 방식의 하나로 제공되고 있다.

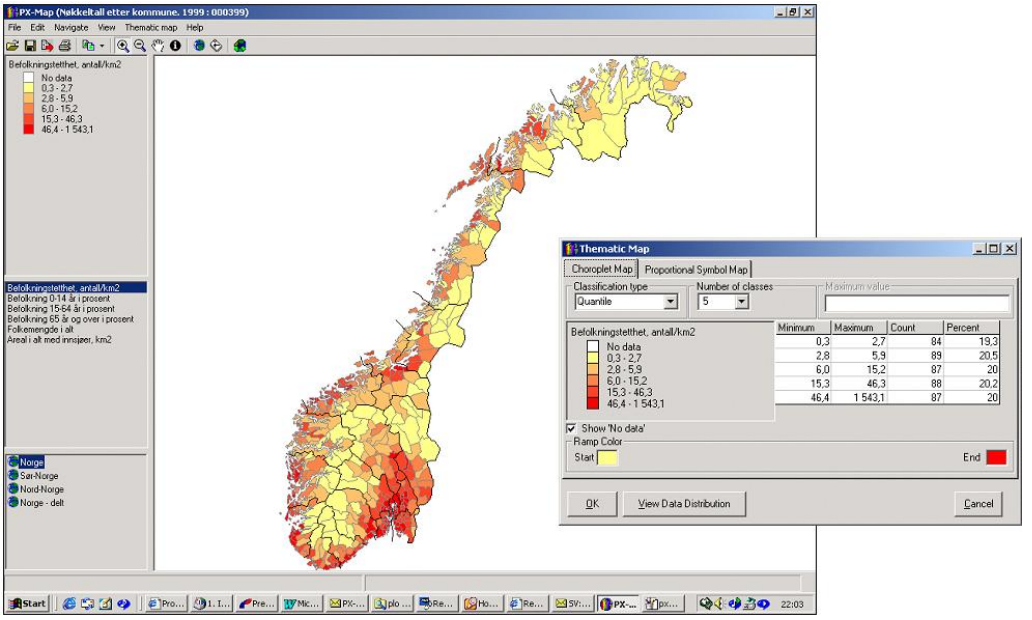
다. 노르웨이

1) 통계지리시스템의 구축

노르웨이의 초기 통계지리시스템 구축은 1993년 통계청과 노르웨이 지도 제작국(Norwegian mapping authority)이 지도 데이터와 통계를 결합한 상품을 개발하기 위한 프로젝트를 위한 실무 그룹을 만들면서 시작되었다. Geodata AS라는 기업과의 파트너쉽을 통해, ArcView, MapObjects를 이용하여 municipality 단위의 통계 주제도를 완성하였다. 이 프로젝트의 주 대상은 중등학교(16-19세) 학생이었으며, 그 외 도서관, 공공 행정기관들도 대상으로 포함되었다. 이 프로젝트는 노르웨이의 인구학적 특성을 배울 수 있는 동적인 학습 툴을 제공하는 데 목적이 있었는데, 지역 통계와 지도를 포함하고 있다.

이후 GIS의 기술적 발전과 함께, 통계의 보다 나은 표현과 해석에 대한 요구, 지역에 대한 지리 참조된 통계에 대한 수요의 증가로 노르웨이 통계청은 GIS를 의제로 선정하고, PC-Axis 소프트웨어 내의 지도 모듈인 PX-Map을 개발하였다(http://www.ssb.no/en/software/pxmap). 이는 통계 주제도를 손쉽게 제작할 수 있는 소프트웨어로, 독립 소프트웨어로도 가능하며, 웹을 통해 무료로 다운로드 받을 수 있다(www.ssb.no/px-map).

PX-Map는 지역 변수들을 포함하는 모든 PC-Axis 파일들의 주제도를 만들 수 있는 소프트웨어로, 대표적인 GIS 소프트웨어 회사인 ESRI의 MapObject에 기반하여 노르웨이 통계청과 Geodata AS에 의해 개발되었다. PC-Map이 지원하는 주제도 유형은 단계구분도, 서클 지도의 두 가지 형태이며, 분위(디폴트)·등간격·사용자 지정의 분류 방법을 제공한다. 계급의 수, 심벌의 크기 및 색상 등을 변경할 수 있으며, 지도의 속성 역시 조작할 수 있는 기능을 가지고 있다. 데이터 분포에 대한 보다 나은 이해를 위한 미리보기 기능(View Data Distribution Function)을 제공한다([그림 6-30] 참조).



[그림 6-30] 노르웨이 PX-Map의 인터페이스와 미리보기 기능

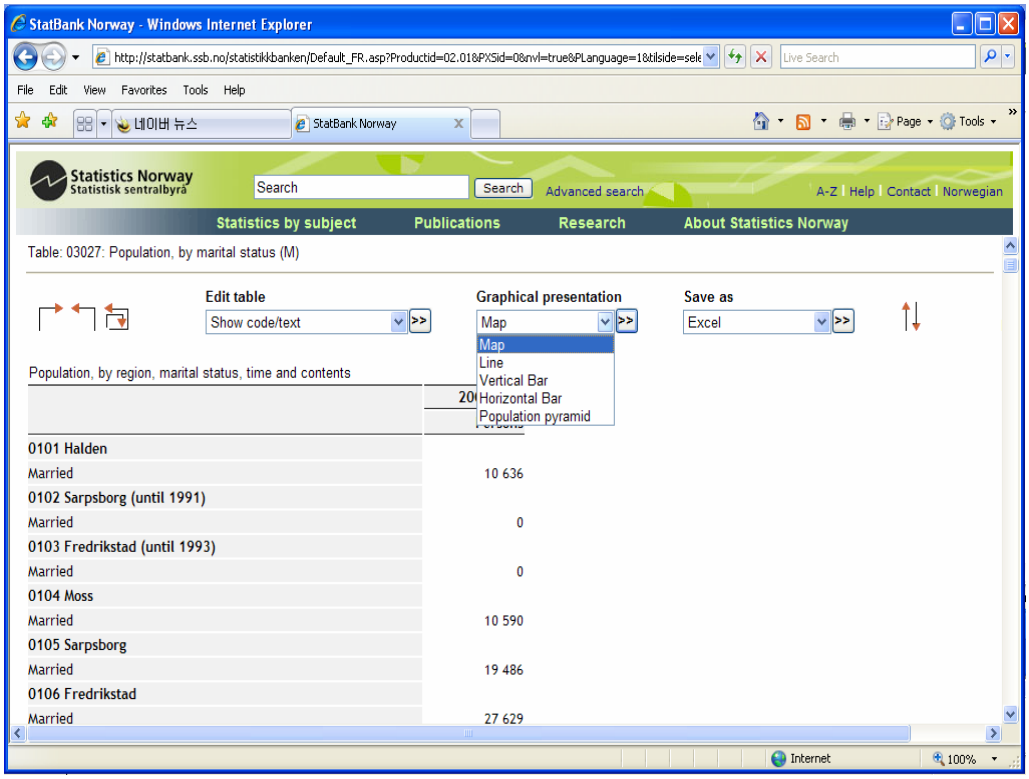
그 외 기능으로, 지도 레이아웃을 그림 포맷(bmp, wmf, png)으로 저장할 수 있다. PX-Map의 개발의 세 가지 원리는 다음과 같다.

- ☐ 상호작용: 사용자와의 상호작용을 극대화시키기 위해 데이터의 분류 방법 및 계급 변경의 옵션을 제공한다.
- ☐ 단순성: 관련 통계와의 결합을 목적으로 하기 때문에 일반 GIS 소프트웨어나 통계 주제도 프로그램에 비해 구조가 단순하다.
- ☐ 오용의 방지: 사용자에 의해 만들어질 수 있는 프로그램과 제한된 옵션들이 시도의 오용을 막는 기능을 한다. 즉, 모든 통계적 변수들이 아니라 지리적 변이를 보이는 속성들만 지도로 표현되도록 제한되며, 통계 속성을 주제도로 표현할 때, 지도학적으로 왜곡된 해석을 가져올 수 있는 오류에 대한 주의 사항이 주어진다. 예컨대, 단계구분도로 총합을 나타내는 것은 자주 범하게 되는 지도학적 오류인데, PX-Map은 이러한 경우, 경고창을 보여줌으로써 오류에 대해 지적하고, 비율 심벌을 이용하도록 한다.

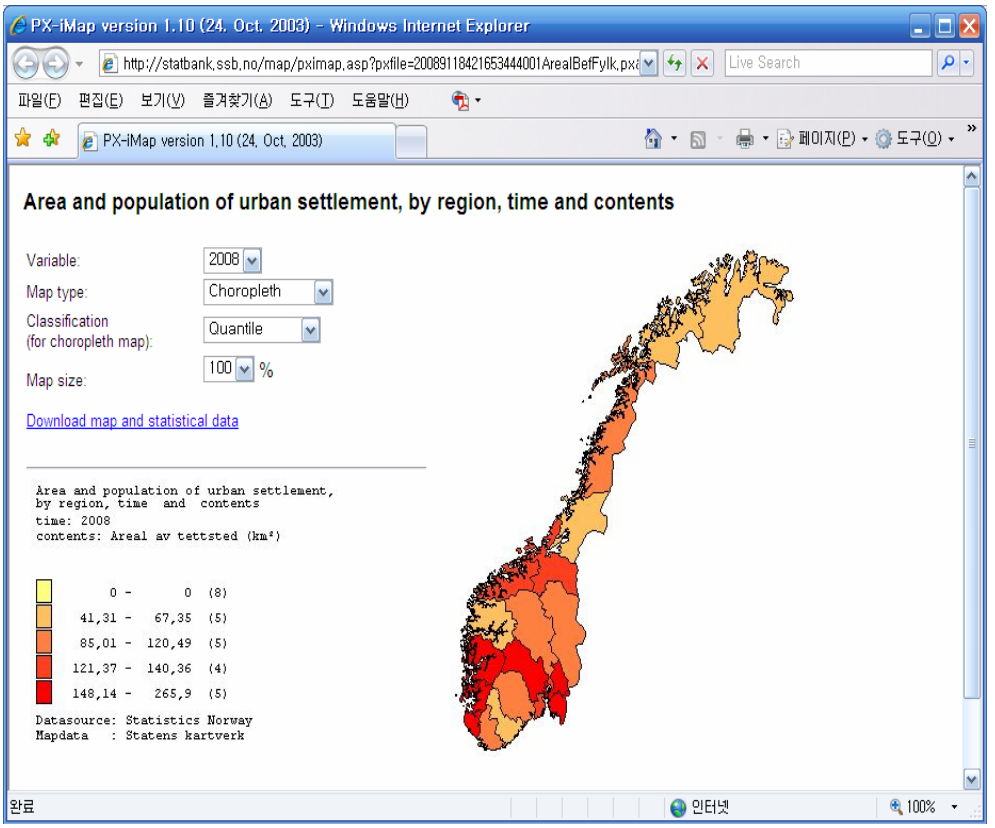
2) 인터넷 GIS 서비스

노르웨이의 인터넷 GIS 서비스는 PX-iMap으로 불리우며, 주제별 속성 데이터, 그래

프, 주제도를 다양한 지리적 단위, 예를 들어, county/municipality/basic statistical unit/urban district/urban settlement 들로 제공하고 있다. 제공되는 주제도의 유형은 단계구분도와 서클 지도이며, 꺾은선 그래프, 막대 그래프, 인구 피라밋과 같은 다양한 그래프의 제공이 특징적이다([그림 6-31]). 분위, 등간격의 데이터 분류법이 제공되지만, 급간의 수, 심벌의 크기 및 색상과 같은 변경이 가능하지 않다는 단점이 있다. 또한 축척과 같은 지도의 주요 요소가 표시되지 않아 지리적 스케일을 파악하기 어려우며, 지도 레이아웃이나 탐색을 위한 브라우징 툴들이 제공되지 않아 사용자 친화적인 인터페이스를 제공하지 못하고 있다. 하지만 인터넷을 통해 통계 및 공간 데이터를 직접 다운로드 받을 수 있는 점은 웹을 통한 통계의 활용 가능성을 높여준다는 측면에서 매우 주목할 만하다. 다운로드 가능한 공간 및 통계 데이터는 PC-Axis와 PX-Map의 고유 파일 포맷으로 제공된다([그림 6-32]).



[그림 6-31] 노르웨이 인터넷 GIS 서비스가 제공하는 다양한 그래프 유형

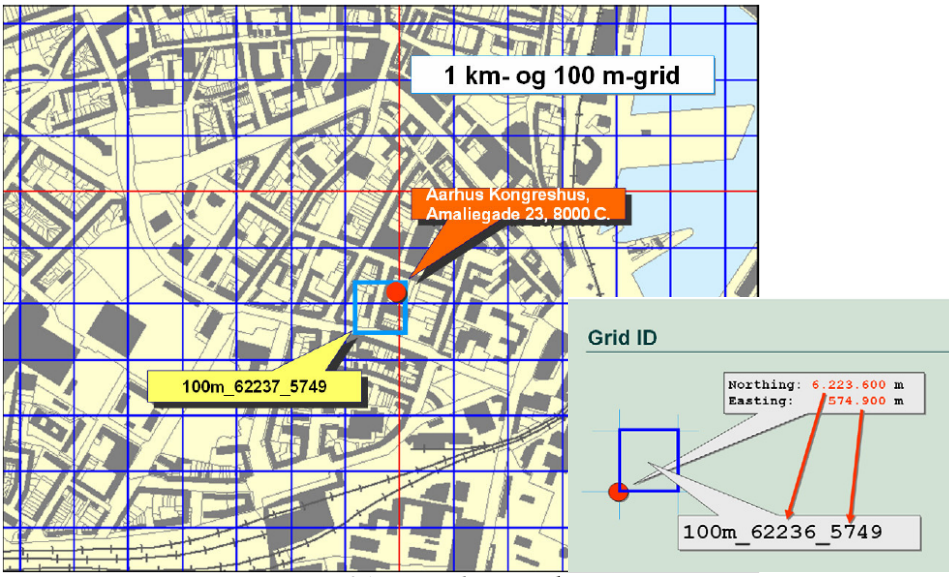


[그림 6-32] 노르웨이 인터넷 GIS 서비스

라. 덴마크

1) 통계지리시스템의 구축

덴마크의 통계지리시스템은 스웨덴, 노르웨이와 같은 등록부 센서스 실시 국가와 동일하게 건물의 좌표의 포인트에 기반하여 구축되었다. Region(기존의 county), municipality, 그리드(1Km, 250m, 100m) 등 통계 속성에 따라 다양한 지리적 단위로 통계가 생산되고 유포되고 있다([그림 6-33]).



Address ID	Municipality-code	Street-code	House-number	Zip-code	Kn100Dk	Kn1kmDk
1810012001_	181	0012	001_	2840	100m_61907_7175	1km_6190_717
1810012001A	181	0012	001A	2840	100m_61907_7175	1km_6190_717
1810012003_	181	0012	003_	2840	100m_61907_7175	1km_6190_717
1810012010_	181	0012	010_	2840	100m_61907_7176	1km_6190_717
1810012011_	181	001				
1810012015_	181	001				
1810032003_	181	003				

Grid ID	Municipality	Households	Population
61901_7126	207	1	6
61902_7126	207	3	8
61903_7126	207	12	26
61904_7126	207	3	7
61905_7126	207	3	8
61909_7126	207	5	10
61910_7126	207	3	6

[그림 6-33] 덴마크의 그리드 참조 시스템

2) 인터넷 GIS 서비스

덴마크의 인터넷 GIS 서비스는 노르웨이의 그것과 매우 유사한데, 동일한 플랫폼인 PX-iMap을 바탕으로 개발되었기 때문이다. 웹 페이지의 외형적 디자인에서부터, 주제도 제작 및 탐색을 위한 인터페이스는 노르웨이와 거의 흡사하다. 하지만 노르웨이의 인터넷 GIS 서비스에 비해 보다 다양한 그래프 형태를 제공한다(예: 파이 그래프 , 막대 꺾은선 혼합 그래프). 한편, 덴마크의 인터넷 GIS 서비스는 통계 및 공간 데이터의 직접 다운로드를 허용하지 않는다.

이상에서 살펴본 북유럽 주요국의 인터넷 GIS 서비스를 국가별 · 항목별로 비교하여 정리한 결과는 다음과 같다(<표 6-6>).

〈표 6-6〉 북유럽 주요국의 인터넷 GIS 서비스 비교

항목	세부항목	스웨덴	노르웨이	덴마크
주제도의 유연성	주제도의 유형	단계구분도, 서클지도, 파이차트 지도, 시계열지도	단계구분도, 서클지도	단계구분도
	분류 방법	등간격, 고정비율	등간격, 분위수	등간격, 분위수
	계급수 변경	○	×	×
	심벌의 크기 및 색상 변경	○	×	×
지도학적 이해	축척, 투영의 표시	×	×	×
사용자 인터페이스	탐색을 위한 브라우징 툴 (확대/축소, 이동, 속성확인)	○	×	×
	출력 레이아웃	×	×	×
공간분석 기능	측정, 공간질의, 버퍼, 레이어 중첩	×	×	×
원자료의 제공	공간 데이터의 직접 제공	×	○	×

제5절 GIS를 통한 소지역 통계의 생산 · 제공과 격자망 데이터

1. 소지역 통계와 격자망 데이터

가. 소지역 통계의 중요성

사회 경제 구조가 복잡해지고 다양해짐에 따라 상세한 지역 통계에 대한 수요는 증

대되고 있으며, 문제 해결과 의사결정을 위한 지식의 창출에 소지역에 대한 센서스 정보는 필수적이다. 소지역에 대한 데이터는 국지화되어 있어서 넓은 지역으로 합역된 데이터에 의해서는 잘 보여지지 않는 사회경제적 현상을 인식하고 이해하는 것을 돕는다(Fiedler et al.). 소지역 단위 공간통계의 필요성과 그 결과의 실질성을 주장하는 많은 연구들이 있다(Dahlberg, 1983; Getis and Ord, 1996; Fotheringham and Brunsdon, 1999; Whitman et al., 2004). 소지역 통계를 유용하게 이용할 수 있는 활용 사례들은 다음과 같다(UN, 2008).

□ 재해재난 구호 계획

산불・지진・화산 등의 자연 재해가 발생한 지역의 인구 및 주택의 특성, 도로・보건소・학교 등과 같은 기반 시설의 유무, 의약품 및 구호 물자의 공급 추정 등에 공간 통계 데이터가 이용될 수 있다. 또한 재해재난의 위험성이 큰 지역에 대한 상세한 평가를 위해서도 소지역 통계를 이용할 수 있다.

□ 공공 서비스 계획

보다 상세화된 인구사회학적 통계 자료를 바탕으로 지역의 공공 서비스의 수요를 예측하고 평가할 수 있으며, 공적 서비스 센터의 자원 분배, 시설물 입지 등 실제적 정책이나 지역 계획과 관련된 다양한 의사결정을 지원할 수 있다.

□ 공적 기반 시설 계획

상수도, 가스, 전기, 통신 등 지역 주민에게 매우 중요한 기반 시설의 제공 및 관리, 미래 수요 예측 등에 활용될 수 있다.

□ 지역 경제 분석

특정 사업체의 고용 및 입지, 접근성, 통근 패턴, 교통, 산업 클러스터 등 지역 경제의 전반적 구조를 파악하거나 지역 계획 및 개발에 필요한 여러 가지 연구들의 수행에 유용한 자료로 활용될 수 있다.

□ 저소득층 지원 프로그램

소지역 공간통계를 바탕으로 저소득층의 분포 및 특성을 파악하고, 주택, 교통, 보건소, 지원센터 등의 기반 시설의 현황을 확인할 수 있으며, 이에 따른 다양한 지역적 지원 프로그램을 수립하는데 이용될 수 있다.

□ 마케팅 분석

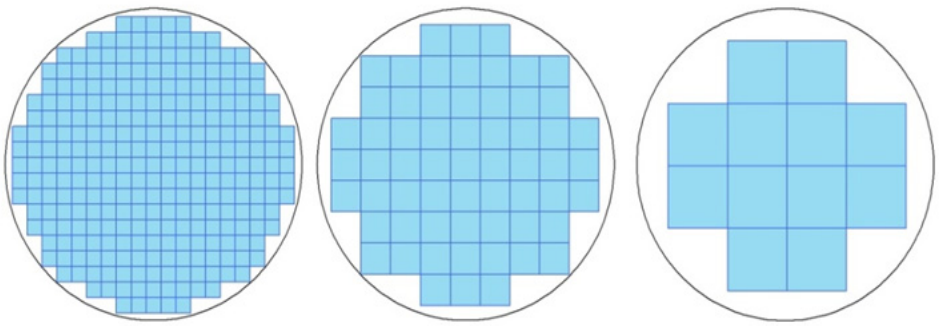
기업의 새로운 점포 입지, 물류 기지의 입지, 특정 홍보대상이나 고객 서비스 정보 관리 등 기업의 마케팅 분석에 활용될 수 있다.

나. 격자망 통계의 특성

소지역 통계의 유용성에도 불구하고 서로 다른 지리적 경계들이 존재하고, 또 그 경계가 지속적으로 변화하는 상황에서 다양한 지리 경계들과 통계 데이터를 통합하는 것은 쉬운 작업이 아니며, 특히 소지역 통계의 생산과 제공에 있어서는 더욱 그러하다. 이러한 상황에 대처하기 위한 하나의 방법은 조사구나 집계구와 같은 작은 공간 단위를 설정한 후 이를 보다 더 큰 공간 단위에 할당하는 것이다. 하지만 이러한 과정은 속성 값의 정확한 맞춤과 안정성을 보장할 수 없는 한계를 가진다. 또 하나의 보다 근본적인 대안은 그리드 참조 시스템(grid reference system)에 기반한 격자망 통계를 제공하는 것이다. 격자망 통계는 통계의 집계 단위를 기존의 행정구역이나 통계 지역이 아닌 경위도에 따른 100m~10km 단위의 그리드 참조 시스템을 사용하는 것으로([그림 6-34]), 데이터의 수집 단위와 집계 단위가 분리되게 된다. 격자망에 의해 지역은 대략 1㎢ 혹은 그 이하의 작은 사각형(mesh)으로 나뉘어진다. 격자망은 지리학에서 분석 수단의 하나로 개발되었는데, 핀란드의 지리학자인 Graneau는 1㎢ 격자망을 사용한 자연·사회 현상의 지역 분석 논문을 발표하였다. 이후 이 방법은 공간 분포 상태와 발전 과정을 분석하고 연구하는데 널리 이용되었다. 격자망 통계를 제공하고 있는 모든 국가에서 기존의 행정구역이나 통계 지역에 대한 통계를 제공하고 있어, 격자망 통계는 기존의 통계를 대체하기보다는 보완하는 특성을 갖는다.

GIS가 제공할 수 있는 데이터 통합 능력, 즉 다른 영역으로부터 다양한 정보들을 연계하는 능력은 격자망 통계를 보다 폭넓게 사용할 수 있도록 하였고, 특히 점점 증가하는 소지역 통계 수요에 대해 각 통계 기관이 즉각적으로 대처할 수 있도록 지원한다. 개별 건물의 정확한 위치가 포인트 데이터로 저장되면, 단순한 GIS 연산, 예를 들어 폴리곤 내 포인트 연산으로 그리드 데이터를 구축할 수 있다.[25]

25) 공간 객체의 위치적 관계를 이용한, 하나의 포인트가 폴리곤 내에 있는지를 판별하는 기하학적 연산이다.



[그림 6-34] 경위도에 따른 그리드 참조 시스템의 설정

격자망 통계는 다음과 같은 장점을 가진다. 첫째, 소지역 통계에 대한 유연하고 안정적인 제공이 가능해진다. 이는 격자망 통계가 가장 낮은 수준의 소스 데이터에 기반하기 때문에, 어떤 형태의 경계 분할에도 정확성을 보장할 수 있기 때문이다. 경계가 변화하는 경우뿐 아니라 어떠한 디지털 지리 경계에 대해서도 그리드 참조 데이터로의 재계산이 가능하다. 예컨데, 행정 구역의 경계가 큰 의미가 없는, 토지이용·자연 등에 의해 정의되는 환경 통계에서, 그리드 참조 시스템은 범람원, 잠재적 오염, 지역별 인구수나 가구수의 재계산을 용이하게 한다. 분할과 합역을 통한 자료 생산이 유연하기 때문에 집계를 통한 새로운 공간 단위의 생성이 용이해 필요로 하는 지역에 관한 데이터를 쉽게 얻을 수 있다.

둘째, 행정경계의 변화와 독립적인 시계열 비교가 가능해진다. 최근 외국의 통계청에서는 행정경계의 변화와 독립적인 데이터 제공 단위에 대한 논의가 증가하고 있으며, 이에 따라 데이터 수집과 독립적인 데이터 제공 단위에 대한 논의 역시 증가하고 있다. 한번 구축된 그리드 시스템은 위치와 분할이 고정되기 때문에 지방 정부의 행정 경계의 변화나 지리적 특성의 변화에 의한 조사구 획정의 변화 등에 영향을 받지 않아 시계열 비교가 용이하다.

셋째, 동일한 그리드 참조 시스템을 통한 자료의 통합이 용이해진다. 특히 통계청 외부의 다른 기관에서 생산되는 자료와의 통합이 용이해지는데, 기상청의 기상 자료와 같이 행정 경계가 의미가 없는 자료와의 통합에 있어 큰 의미를 가질 것이라 예상된다.

넷째, 그리드 참조 시스템은 GIS의 기능과 연결되어 다양한 형태의 분석에 이용될 수 있다. 예컨데, 어떤 재해의 발생 지점이나 특정 교통 경로에서 정해진 거리 이내에 있는 세대를 확인하거나, 공간 내삽을 통해 결측된 소지역에 대한 데이터 값을 추정할 수도 있다.

다섯째, 거의 크기와 형태가 동일하게 분할되기 때문에 격자망 간 특성을 측정하고

비교하는 것이 용이하다. 격자망은 거의 경도와 위도에 의해 네모 모양으로 분할되기 때문에 위치가 쉽고 분명하게 인식될 수 있으며, 거리에 관한 데이터를 쉽게 분석하고 계산하고 비교할 수 있다.

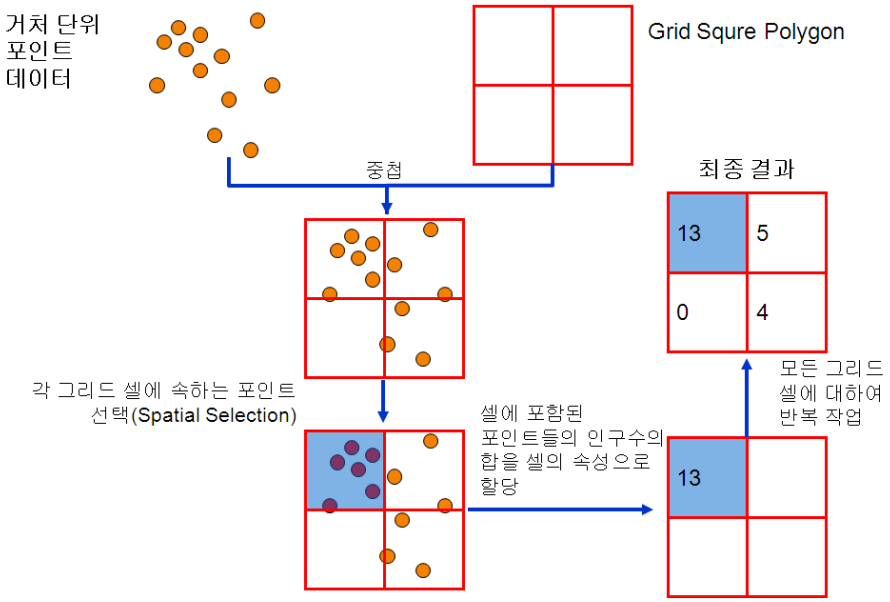
많은 장점에도 불구하고 격자망 통계를 제공하고 있는 모든 국가에서 기존의 행정구역이나 통계 지역에 대한 통계를 제공하고 있어, 격자망 통계는 기존의 통계를 대체하기보다는 보완하는 특성을 갖는다.

다. 격자망 통계의 생산 방법과 우리나라 통계청의 자료 구축 현황

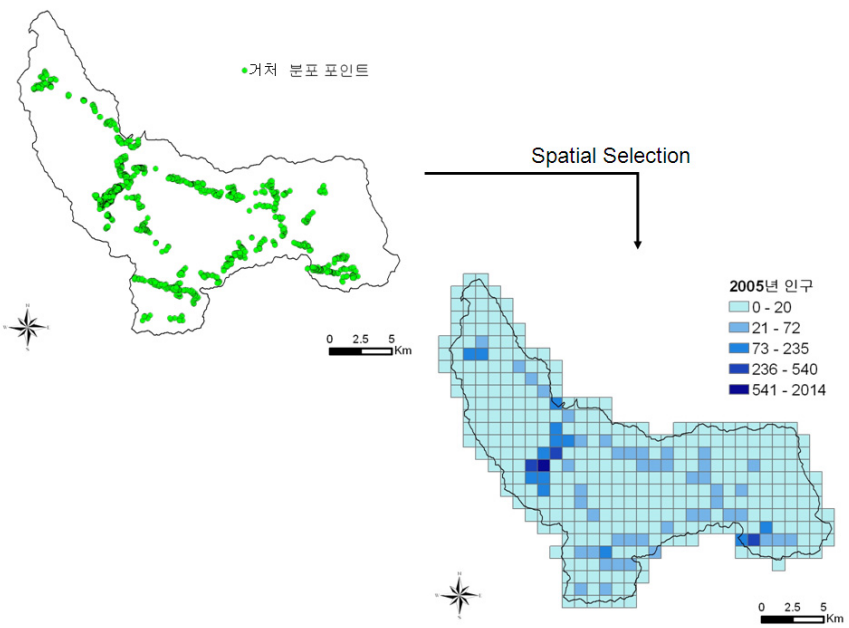
격자망 통계는 기반이 되는 데이터의 유형과 이에 따른 통계 생산 방법에 따라 포인트 기반 격자망 통계와 폴리곤 기반 격자망 통계 두 가지로 구분될 수 있다. 우선 포인트 데이터에 기반해서 GIS의 공간 선택(spatial selection) 기능에 의해 격자망 통계를 생산하는 과정을 살펴보고, 그 결과를 산출해 본 결과는 다음과 같다([그림 6-35], [그림 6-36]). 포인트 단위의 데이터에 그리드 셀을 중첩(overlay) 시킨 후 각 그리드 셀에 속하는 포인트를 공간 선택에 의해 선택하게 된다. 각 셀에 포함된 포인트들의 인구수를 합하여 각 셀에 할당하면 격자망 통계가 만들어진다. 위와 같은 폴리곤 내 포인트 연산은 상용 GIS소프트웨어에 기본 기능으로 포함되어 있다.

폴리곤 기반 데이터를 또 다른 형태의 폴리곤인 격자망 통계로 재집계 하는 데는 기술적인 어려움이 있다. 이는 격자망이 블록이나 조사구와 같이 집계되는 소지역과 일치하지 않기 때문이다. 격자망은 경도와 위도에 의해 경계가 지워지는 반면 집계 소지역들은 도로, 강, 철도 등과 같은 지리적 지형・지물에 의해 경계가 지워진다. 따라서 데이터의 정확성을 유지하기 위해 폴리곤 데이터의 격자망을 편집하는 일은 시간과 노동력이 많이 투여되어야 하는 일이다. 일본 통계국은 격자망의 편집 작업을 향상시키기 위해 오랫동안 다양한 대응책을 취해왔다(일본 통계국 홈페이지).

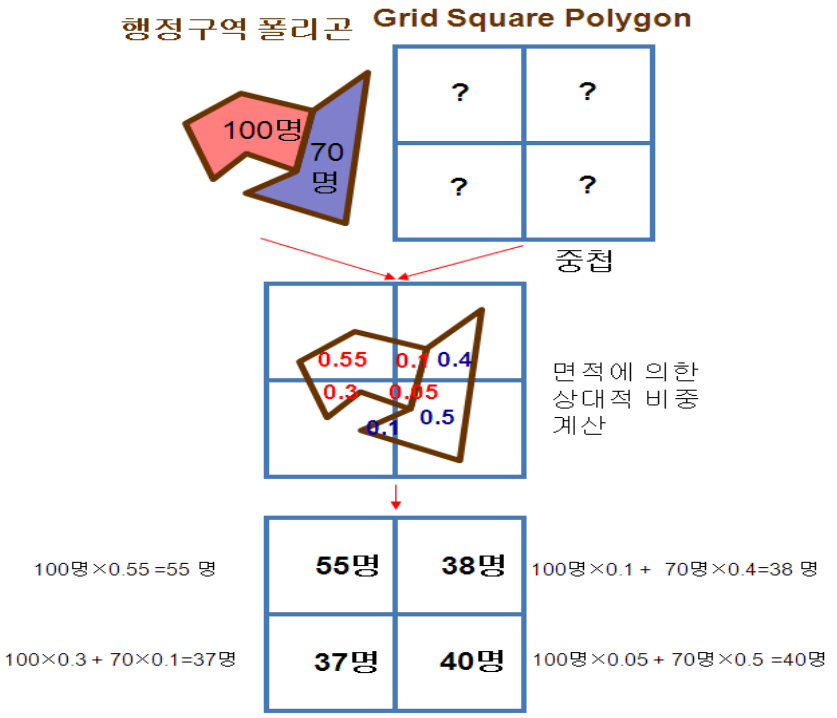
폴리곤 데이터에 기반해서 격자망 통계를 생산하는데 있어서도 GIS의 면 내삽(areal interpolation) 기능은 중요한 역할을 할 수 있을 것이라 예상된다. 하지만 면 내삽은 폴리곤내 포인트 연산과 달리 아직 상용 GIS소프트웨어에 기본 기능으로 포함되어 있지 않다. 이에 폴리곤 기반 데이터의 내삽법에 의한 격자망의 인구수 산출 과정을 개념적으로 살펴본 후([그림 6-37]), 프로그래밍을 통해 그 결과를 산출해 본 결과는 다음과 같다([그림 6-38]). 폴리곤 기반 데이터는 포인트 기반 데이터와 같이 개별 통계값을 가지는 것이 아니라 이미 집계된 통계값을 가지고 있기 때문에, 단순 합계가 아닌 상대적 비중을 계산하였다. 우선 행정구역 폴리곤과 그리드 셀을 중첩시킨 후 각 행정구역의 그리드 셀에 대한 상대적 비율을 계산하였다. 계산된 면적 비율에 따라 행정구역의 인구를 개별 그리드 셀에 할당함으로써 격자망 통계를 생성하였다.



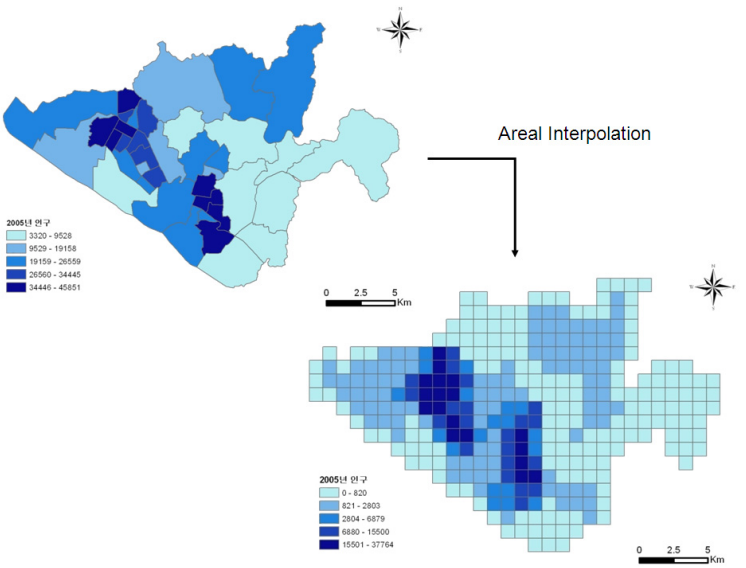
[그림 6-35] 포인트 기반 데이터의 공간 선택에 의한 격자망의 인구수 산출 과정



[그림 6-36] 포인트 기반 데이터의 공간 선택에 의한 격자망의 인구수 산출 결과

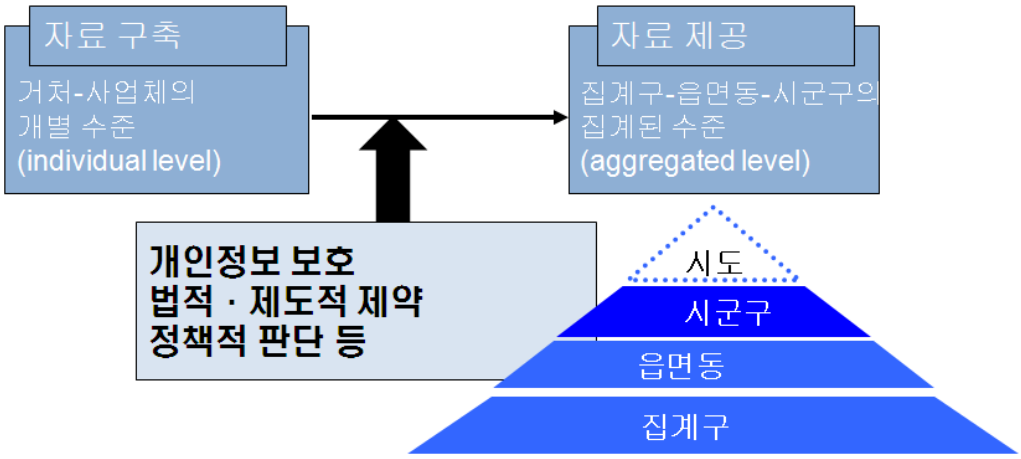


[그림 6-37] 폴리곤 기반 데이터의 면 내삽법에 의한 격자망의 인구수 산출 과정



[그림 6-38] 폴리곤 기반 데이터의 면 내삽법에 의한 격자망의 인구수 산출 결과

우리나라의 통계청은 전국 약 1,350만 개의 거처와 320만 개 사업장에 대한 위치 정보와 통계 정보를 결합한 포인트 데이터 기반의 공간 데이터베이스를 구축하고 있다([그림 6-39]). 개인정보 보호, 법적·제도적 제약, 정책적 판단 등에 따라 현재까지는 집계구-읍면동-시군구-시도 단위의 집계 데이터가 제공되고 있다. 향후 통계청의 위치 정보를 가지고 있는 거처 포인트 데이터는 GIS의 폴리곤 내 포인트 기능을 통해 격자 통계로 가공될 수 있어, 격자 통계 제공의 기반이 될 수 있을 것이라 예상된다.



[그림 6-39] 우리나라 통계청의 자료 구축과 제공 현황

2. 등록부 기반 센서스 실시 국가와 일본의 격자망 데이터

가. 포인트 데이터 기반의 격자망 데이터: 등록부 기반 센서스 실시 국가

핀란드·노르웨이·스웨덴 통계청은 공동으로 Tammilehto-Luode et al.(2000)의 보고서 'GRID 데이터와 정의에 따른 권역 설정: 보다 나은 유럽 지역 통계 시스템을 위해'를 제출하였다. 이를 중심으로 등록부 기반 센서스 실시 국가의 그리드 데이터 제공 현황을 살펴본 결과는 다음과 같다.

EU의 지역 통계는 NUTS(Nomenclature des Unites Territoriales Statistiques)에 기반하여 관리되어 왔다. 이는 EU의 지역 통계 생산에 있어 단일하고 동질적인 지역 구분을 제공하기 위해 Eurostat에 의해 만들어졌다. 그러나 많은 분석에 있어 NUTS는 이상적이지 않다. 한 레벨에 속하는 단위들 간에 크기의 차가 존재하고, 시간에 따라 단위가 변해왔기 때문이다. NUTS의 지역은 일반적으로 국가의 행정 단위이다. 하지만 많은 경우 행정적

경계는 자연 현상이나 혹은 관심 대상이 되는 현상에 의해 구분되는 지역과 일치하지 않는다.

등록부에 기반하여 센서스를 실시하고 있는 북유럽 국가들은 유럽 전역 특히 Eurostat의 그리드 형태의 통계 사용을 촉진시켜왔다. 유럽 대부분의 국가 통계 기관들은 통계 생산에 있어 그리드 시스템을 이용하고 있다(Tammilehto-Luode et al, 2000). 등록센서스가 실시되고 있는 핀란드・노르웨이・스웨덴의 통계 단위의 지오코딩은 포인트 기반이다. 핀란드에서 지오코딩의 1차적 소스는 건물들의 중심점이고, 스웨덴에서는 부동산의 중심점이며, 노르웨이에서는 표준화된 주소들이다. 지오코딩된 점들과 통계 단위와의 연결은 소규모 지역 통계 생산을 유연하게 만들어 준다. 이것은 왜 상대적으로 작은 지역들, 1kmx1km 그리드 사각형 또는 보다 작은 단위의 통계가 생산되고 있는지, 스웨덴과 핀란드에서 소지역 데이터가 광범위하게 사용되고 있고, 노르웨이에서도 소지역 데이터의 사용이 증가되고 있는지에 대한 주요 이유일 것이다. 세 나라는 인구의 공간 구조를 비교・대조하기 위해 그리드 셀을 사용하여 실험하였다.

나. 폴리곤 데이터 기반의 격자망 데이터: 일본[26)]

1) 개관 및 격자망 통계의 역사적 배경

소지역 통계에 대한 요구에 부응하기 위해 일본 통계국에서는 '격자망 기법(grid square method)'이라는 명칭으로 격자망 통계를 제공하고 있다. 통계국은 인구 센서스 결과와 사업체 센서스 결과를 경・위도에 의해 일본의 전지역을 포괄하는 대략 1㎢와 0.25㎢의 작은 격자별로 집계할 수 있도록 재구조화하고 있다. 격자망 통계는 지역 개발, 도시 계획, 교통 계획, 오염 자연 재난에 대한 대책을 포함하는 여러 다양한 행정적 대응을 위한 계획을 세우기 위한 단순한 숫자값일 뿐 아니라 지역에 대한 배경 정보가 된다.

격자망 통계는 일본 전역을 위선과 경선에 기반해서 대략 1㎢(1㎞ 간격)와 0.25㎢(500m 간격)의 작은 사각형(mesh)으로 나뉘어진 가장 소지역에 관한 통계이다. 격자망 통계는 지역별 실제 상황을 같은 기준하에서 보다 자세하게 지역별로 보게 해주기 때문에, 공공 영역과 사적 영역 모두에서 여러 가지 목적으로 널리 사용된다. 국가 정부와 지역정부에서는 도시 계획, 지역 개발, 재해 예방과 환경 프로그램, 환경 오염 대책에, 민간 부문에서는 시장과 상업 지역 분석, 학문 연구 등에 사용된다.

일본 통계국에서 국가 통계에 이 방법을 사용하기 시작한 것은 경제가 급격하게 성장하던 1969년이다. 1969년에 총무성(당시는 Administrative Management Agency이며, 현재는 Ministry of Internal Affairs and Communications) 산하의 통계국에서는 시험적으로

26) 일본의 격자망 데이터는 일본 통계국의 홈페이지의 내용을 중심으로 살펴보았다.

1965년 인구 센서스, 1966년 사업체 센서스, 1968년 주택 센서스에 기반해서 동경 대도시권과 같은 특정 지역에 대해 '국토의 실제 상황에 관한 통합된 통계(Intergrated Statistics for the Actual Situation of National Land)'라는 제목으로 격자망 통계를 작성하였다. 이후 정부 부문뿐만 아니라 사적 영역에서도 격자망 기법을 사용한 지역 정보의 수집과 관리가 촉진되었다.

격자망 통계가 중요해진 이유는 소지역 정보(small area information)에 대한 수요의 증가에 있다. 우선 1950년대에 시작된 경제 성장은 도시 지역으로의 인구 집중과 함께 급격한 사회경제적 변화를 야기하였다. 균형 성장을 실현하기 위해, 미래의 지역 격차를 줄이는 것은 국가 정부와 도 단위뿐 아니라 기초 자치단체에서도 필수적이 되었다. 이와 함께 사적 영역에서는 시장 조사와 합리적인 경영을 위해 어떤 일이 발생하고 있고 어떤 일이 발생할 것인지를 알기 위한 기본 정보로서 보다 상세한 지역 정보를 원하게 되었다.

둘째, 1953년에 실시된 행정구역 통합 촉진 법률(Town and Village Merger Acceleration Law)에 근거해서 많은 작은 지역들이 통합되었다. 결과적으로 일본의 자치단체(municipals)는 1950년 인구 센서스 당시 10,414개에서 통합이 완료된 1960년에는 3,511개로 감소하였다. 동경은 23개의 특별구(ward)에서 하나의 시로 되었다. 결과적으로 지방자치단체의 평균 인구 규모와 면적은 각각 1950년에 인구 8,000명 36㎢에서 1960년 27,000명, 107㎢로 세 배 가까이 증가되었다. 그러나 통계 데이터는 여전히 자치단체 별로 집계되었고, 따라서 더 작은 지역에 대한 통계 자료에 대한 수요가 증가하였다.

셋째, 당시 지방자치단체들은 수도, 하수도와 같은 공공 시설의 건설을 시작하였는데, 자치단체별로 집계된 통계 자료는 사회 인프라의 향상과 같은 공공 계획의 설정에 불충분하게 되었다.

이러한 문제를 해결하고 다양한 소지역의 통계를 수집하고 비교하기 위해 격자망 기법의 채택이 제안되었다. 이후 현재의 '격자망'이 많은 토론과 고려 이후에 설계되었다. 1973년 총무성은 통계 목적의 표준 격자망과 격자망 코드를 발표하였다(1973년 7월 12일 훈령 143호, <부록 4> 참조).

2) 격자망의 획정 방법

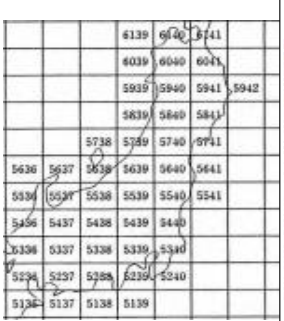
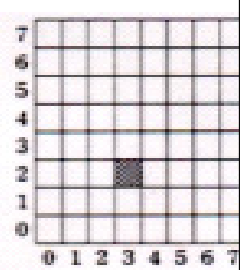
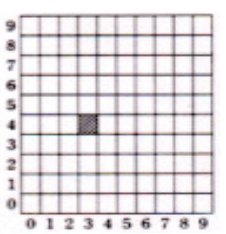
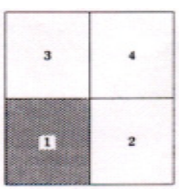
표준 격자망의 획정 방법은 다음과 같다(<표 6-7>, <표 6-8>). 1차 지역 구획(Primary Area Partition)은 위도 40분, 경도 1도에 의해 나뉘어진다([그림 6-40]). 2차 지역 구획(Secondary Area Partition)은 1차 구획에 의해 나뉘어진 지역을 가로 세로 8등분하여(8X8) 64개의 동일한 크기의 지역으로 나눈다. 3차 지역 구획(Third Area Partition)에 의한 기본 격자망(Basic Grid Square)은 2차 구획에 의해 나뉘어진 지역을 가로 세로 10등분하여(10X10) 100개의 동일한 크기의 지역으로 나눈다.

〈표 6-7〉 격자망 경계의 획정 방법

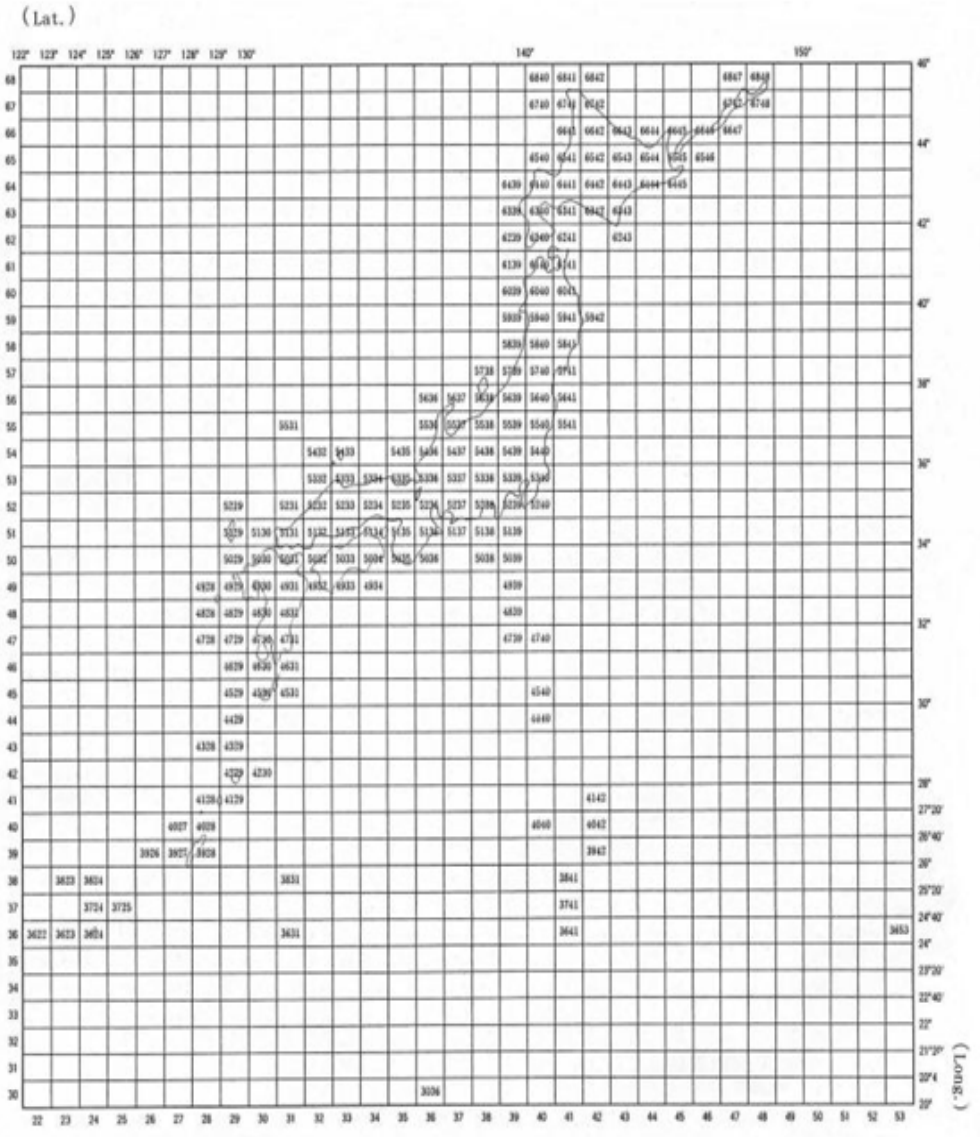
구획	경계를 나누는 방법	위도 간격	경도 간격	길이	지도와의 관계
1차 구획	국가 전체를 위도 40분과 경도 1도로 나눔	40분	1 도	약 80 km	1/200,000 지형도(Geographical Survey Institute)
2차 구획	1차 구획 지역을 동일한 크기의 64(8X8)개의 지역 으로 나눔	5 분	7 분 30 초	약 10 km	1/25,000 지형도(published by the Geographical Survey Institute)
3차 구획 (기본 격자)	2차 구획 지역을 동일한 크기의 100(10X10)개의 지역 으로 나눔	30 초	45 초	약1 km	-
분할된 2분 격자망	기본 격자를 동일한 크기의 4(2X2)개의 지역 으로 나눔	15초	22.5 초	약 500 m	-

http://www.stat.go.jp/english/data/mesh/05.htm.

〈표 6-8〉 격자망 경계의 획정 결과

1차 구획	2차 구획	3차 구획 (기본 격자)	분할된 격자망 (2분 격자망)
동경 대도시권을 포함하는 1차 구획	2차 구획 결과	3차 구획 결과	기본 격자의 분할 결과
코드: 5339	5339-23	5339-23-43	5339-23-43-1
			

http://www.stat.go.jp/english/data/mesh/05-1.htm.

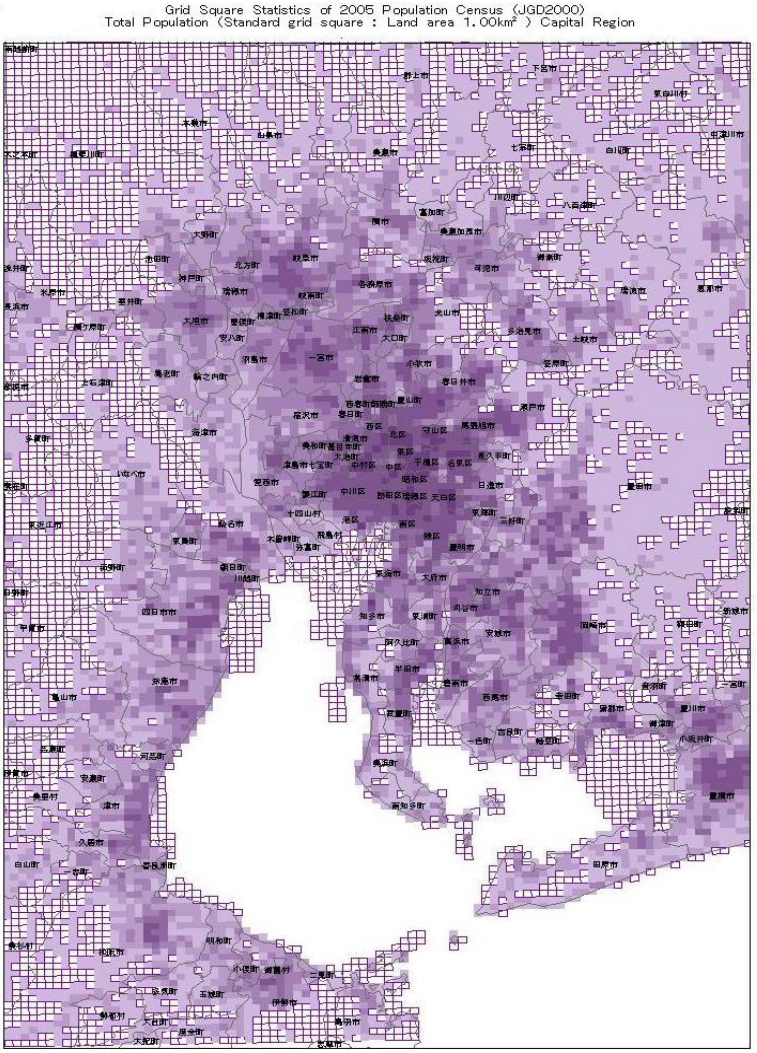


출처: http://www.stat.go.jp/english/data/mesh/

[그림 6-40] 일본 전국의 1차 구획에 의한 격자망

3) 일본 통계국에 의해 생산되는 격자망 통계

일본 통계국에 의해 생산되는 격자망 통계는 다음과 같다(<부록> 참조). 일본 통계국의 격자망 통계 제공 사례를 2005년 인구 센서스의 동경 대도시권 총인구수의 1㎢ 격자망 데이터를 통해 살펴본 결과는 다음과 같다([그림 6-41]).



출처: http://www.stat.go.jp/english/data/mesh/

[그림 6-41] 2005년 인구 센서스 동경 대도시 총인구의 1㎢ 격자망 통계

□ 2005년 인구 센서스의 격자망 통계
인구 센서스에 기반한 인구, 가구 등에 관한 격자망 통계로, 연령별 직업별 건물 유형별 가구 등을 포함한다.

□ 2006년 사업체와 기업(establishment and enterprise) 센서스의 격자망 통계
사업체·기업 센서스에 기반한 사업체, 고용 등에 관한 통계로, 산업별 고용자수, 고용 규모별 기업 등의 숫자를 포함한다.

□ 2000년 인구 센서스와 2001년 사업체와 기업 센서스가 결합된 격자망 통계
이 격자망 통계는 인구센서스와 사업체·기업 센서스가 결합되어 있는데, 이 통계는 주간 인구와 인구당 사업체의 개수 등에 관한 것을 포함한다.

제6절 결 론

본 연구는 문헌 연구를 통해 공간통계 생산과 제공에 있어서의 GIS의 역할에 관해 이론적·논리적으로 살펴보고자 하였다. 이를 위해 우선 미국 센서스국 지리실(Geography Division)이 데이터의 수집·처리·집계를 향상시키는 혁신적인 방법을 고안하는 과정에서 GIS의 발달에 어떻게 공헌하였는지를 살펴보았으며, 다음으로 센서스 조사 방법이 다른 현장 조사 실시 국가와 등록센서스 실시 국가에서의 센서스 GIS에 대해 살펴보았다. 이와 함께 GIS에 기반한 소지역 통계 제공의 일환으로 격자망 통계의 제공 가능성을 살펴보고, 격자망 통계 생산 방법을 검토해 보았다. 본 연구의 결과는 다음과 같다.

미국의 센서스국, 특히 지리적·지도학적 측면에서 센서스 업무를 지원하는 지리실은 GIS의 역사에서 중요한 역할을 담당하였다. 조사 방법에 따라 지리실의 역할이 변하였지만, 핵심 역할은 거주지의 지리적 위치를 결정(geocoding)하고, 거주지를 집계 영역으로 분류(tabulation)하는 것이었다. 우편에 의한 센서스의 실시에 의해서 지오코딩 과정을 자동화할 필요성이 제기되었고, 이를 위한 다양한 방법이 고안되었다. 자동화된 지오코딩은 수치화된 공간정보를 필요로 하기 때문에 센서스국에서는 도로, 센서스 경계, 행정구역 경계, 기타 공간 형상을 수치화하기 시작하였다. 센서스국의 노력의 산물로서 ACG, DIME, GBF/DIME, 타이거 등이 등장하게 되었다. 이 과정에서 미국 센서국이 GIS의 발달에 미친 영향을 살펴본 결과는 다음과 같다. 첫째, GIS 데이터베이스에서 위상구조의 개념을 정립하고 보편화시키는 역할을 하였다. 이는 수치 환경에서 공간데이터를 관리하고 공간 객체 사이의 관계를 정의하는데 있어 매우 가치 있는 성과라 할 수 있

다. 둘째, 타이거 데이터베이스의 구축은 국가적 규모의 공간 데이터베이스를 만들어 냈다는 점에서 의의가 있다. 타이거의 구축을 통하여 국가적 공간분석이나 의사결정이 가능하게 되었다. 셋째, 타이거 데이터의 무료 배포는 GIS의 보편화와 대중화에 기여하였다. 넷째, MAF/TIGER 강화 프로그램을 통하여 구축된 향상된 공간데이터는 향후 국가 공간 데이터 인프라(NSDI) 역할을 담당할 것으로 기대된다. 타이거의 일차적인 목적은 센서스의 지원이지만, 향후 수많은 잠재적 사용자의 요구를 만족시키기 위한 고려가 이루어지고 있다.

개별 국가가 채택하고 있는 센서스 조사 방법, 즉 전통적인 현장 조사 방식인지 등록부에 기반한 방식인지에 따라 개별 국가들은 서로 다른 형태의 GIS를 구축하고 있다. 하지만 센서스의 조사 방법과 상관없이 국가 통계 기관에서는 보다 나은 통계의 생산과 제공을 위해 통계와 GIS를 결합하기 위한 노력을 지속하고 있다. 조사 기반 센서스 실시 국가에서는 센서스 전 단계에서 GIS를 활용하고 있는데, 이를 단계별로 나누어 살펴보면 다음과 같다. 센서스 전단계에서는 통계 단위의 설정과 관리, 조사구의 설정 및 관리, 조사구요도의 제공에 활용된다. 센서스 단계에서는 현장 조사의 관리・업데이트・새로운 조사방법의 도입에 활용되며, 센서스 사후 단계에서는 집계 통계의 생산・센서스 결과의 주제도화・인터넷 GIS를 통한 다양한 형태의 지도 출판에 사용된다.

현장 조사 없이 등록부에 기반하여 센서스를 실시하고 국가의 센서스는 포인트 객체를 바탕으로 한 GIS에 기반하고 있다. 개별 주소, 빌딩과 같은 포인트를 최소 입력 단위로 하는 공간 데이터베이스 구축을 통해 다양한 공간 단위별로 유연하게 데이터를 생산・제공하고 있다. 한편 북유럽 국가들은 통계 및 공간 데이터에 대한 접근을 위한 무료 소프트웨어를 자체적으로 제작하여 사용하고 있다. 스웨덴에서 개발된 PC-Axis 프로그램과 그 외 북유럽 국가들에서 개별적으로 개발된 하위 모듈들은 북유럽 지역 통계 생산 및 배포를 위한 호환가능한 툴을 제공함으로써, 통계 데이터의 표준화와 비교 분석에 중요한 기여를 하고 있다.

출력된 보고서, CD/DVD와 같은 저장 미디어를 통해 배포되던 각종 통계 데이터는 오늘날 대부분의 국가에서 인터넷을 통해 제공되고 있다. 특히 인터넷 GIS를 통한 공간 통계 제공이 이루어지고 있는데, 이는 센서스의 조사 방법과 상관없이 많은 국가에서 이루어지고 있다. 일반적으로 지도는 통계 데이터베이스와 연동되어 테이블, 그래프와 함께 통계 데이터를 표현하는 또 다른 시각적 자원으로 제공되고 있다. 인터넷 GIS는 각국에 따라 상이한 수준의 서비스와 인터페이스를 갖추고 있는데, 대체적으로 주제도의 형태, 계급분류 방법, 심벌의 변경 등 주제도 제작에 있어 사용자에게 유연성을 제공하려고 노력하고 있다.

1990년대 이후 GIS의 발전과 컴퓨터 기반 속성 자료 입력과 경계선 입력・획정 기법

의 발전으로 인해 특수한 목적에 맞게 센서스 공표구역을 획정할 수 있게 되었다. 이에 최근 외국의 국가 통계 기관에서는 행정경계의 변화와 독립적인 데이터 제공 단위에 대한 논의가 증가하고 있으며, 이에 따라 데이터 수집과 독립적인 데이터 제공 단위에 대한 논의 역시 증가하고 있다. 영국 통계청은 2001년 센서스에서 GIS에 기반해서 새로운 공표구역인 Output Area를 생성함으로써 조사구와 공표구역을 분리하였다. 한편, 사회경제 구조가 복잡해지고 다양해짐에 따라 문제 해결과 의사결정을 위한 지식의 창출과 관련 소지역 통계에 대한 수요가 증대되고 있다. 일본에서는 소지역 통계에 대한 수요 증가와 행정구역 통합에 따른 대응으로서 격자망 통계가 1969년부터 제공되고 있다. 격자망 통계는 행정경계의 변화와 독립적인 시계열 비교가 가능하고, 동일한 그리드 참조 시스템을 통한 자료 통합이 용이해지는 등 여러 장점을 가지는데, 이에 GIS에 기반한 소지역 통계 제공의 일환으로 격자망 통계의 제공 가능성과 격자망 통계 생산 방법을 검토한 결과는 다음과 같다. 우리나라 통계청은 전국 약 1,350만 개의 거처와 320만 개 사업장에 대한 위치 정보와 통계 정보를 결합한 포인트 데이터 기반의 공간 데이터베이스를 구축하고 있다. 현재까지는 집계구-읍면동-시군구-시도 단위의 집계 데이터가 제공되고 있지만, 향후 통계청의 거처 포인트 데이터는 GIS의 폴리곤 내 포인트 연산 기능을 통해 격자 통계 제공의 기반이 될 수 있을 것이라 예상된다.

참고문헌

강영옥・윤은주・원종석・이성호・정재희, 2007, 2006년도 국가 GIS 지원 연구: GIS 기반 소지역 통계집계 공표구역의 획정 및 관리방안 연구.

강영옥・장세진・박보람, 2008, 2007년도 국가GIS 지원연구: 행정 및 통계 자료 축적을 위한 표준 통계 단위 설정 연구.

건설교통부, 2007, GIS 기반 소지역 통계집계 공표구역의 획정 및 관리방안 연구, 2006년도 국가 GIS 지원 연구.

국토해양부, 2008, 국토지리정보원 발전방안 수립 연구.

김형석・양경진・강영민, 2003, 인구센서스와 행정자료의 활용, 한국조사연구학회 춘계학술대회 발표문.

변필성・김광일・안재학, 2006, 소지역 통계 공표구역의 사례: 잉글랜드・웨일즈의 Output Area, 국토, 제300호, 92-99.

신정엽・이상일, 2008, GIS의 개념과 원리, 다락방.

이상일・신정엽・김현미・홍일영・김감영・전용완・조대헌・김종근・이건학(역), 2009, 지리정보시스템과 지리정보과학, 시그마프레스. Longley, P. A., Goodchild, M. F., Maguire, D. J. and Rhind, D. W., Geographic Information Systems and Science, 2005, John Willy & Sons Ltd.

조순기・최은영, 2009, 행정자료 기반의 등록센서스 해외 사례 연구, 통계개발원.

통계청, 2008a, 공간통계지식체계 기반표준 개발사업.

통계청, 2008b, 통계지리정보 중장기 발전계획 수립.

통계청, 2007, 공간통계 정보체계 구축 및 활용촉진을 위한 법제도입방안 연구.

Amar, D. and Shahar, K., 2001, GIS and the population and household census in Israel from present to future, Proceedings of ESRI International User Conference, July 9-13, San Diego, USA.

Beddoe, D., 1999, Census 2000 count: some GIS ideas, ArcNews, ESRI, http://www.esri.com/news/arcnews/fall99articles/38-census2000.html.

Benenson, I. and Omer, I., 2003, High-resolution census data: simple ways to make them useful, Data Science Journal, 2, 117-127.

Booz-Allen and Hamilton., 2000, Master Address File(MAF)/TIGER Modernization Study, http://www.census.gov/geo/mod/booze_zllen.pdf.

Broome, F. R. and Meixler, D. B., 1990, The TIGER Data Base Structure, Cartography and Geographic Information Science, 17(1), 39-47.

Carbaugh, L. W. and Marx, R. W., 1990, The TIGER System - A Census Bureau innovation serving data analysts, Government Information Quarterly, 7(3), 285-306.

Cooke, D. F., 1998, Topology and TIGER: The Census Bureau's contribution, in The History of GIS: Perspectives from the Pioneers, T. Foresman(eds.), Prentice Hall, New Jersey, 47-57.

Coppock, J. T. and Rhind, D., 1991, The history of GIS, in Maguire, D. J., Goodchild, M. F., and

Rhind, D.W.(eds.), Geographical Information Systems, Volume 1: Principles, Longman, Harlow, Essex, 21 – 43.

Dahlberg, K. A., 1983, Contextual analysis: taking space, time, and place seriously, International Studies Quarterly, 27(3), 257-266.

ESRI, 2006, US census 2010 supported by ArcPad mobile GIS technology, ArcNews, ESRI.

Fiedler, R., Schuurman, N., Hyndman, J., 2006, Improving census-based socioeconomic GIS for public policy: recent immigrants, spatially concentrated poverty and housing need in Vancouver, An Introductional E-Journal for Critical Geographies, 4(1), 145-171.

Fotheringham, A. S. and Brunsdon, C., 1999, Local forms of spatial analysis, Geographical Analysis, 31(4), 340-358.

Getis, A. and Ord, J. K., 1996, Local spatial statistics: an overview, in Longley, P. and Batty, M.(eds.), Spatial Analysis: Modelling in a GIS Environment, GeoInformation International, Cambridge, 261-277.

Goodwin, A. and Daniel, C., 2002, Maximizing use of census 2000 data in local government, Proceedings of ESRI International User Conference, July 7-12, San Diego, USA.

Gould, M. and Hecht, L., 2001, A framework for geospatial and statistical information integration, Working paper No. 27, Joint in UNECE/EUROSTAT Work session on Methodological Issues Involving the Integration of Statistics and Geography, Conference of European Statisticians, Tallinn, Estonia.

Jamagne, P., 1998, Step towards GIS development at the National Statistical Institute of Belgium, Working paper No. 20, Work Session in Geographical Information Systems, Conference of European Statisticians, Ottawa, Canada.

Laurila. S., 2001, Serving customers with GIS services, Working paper No. 33, Joint UNECE/EUROSTAT Work session on Methodological Issues Involving the Integration of Statistics and Geography, Conference of European Statisticians, Tallinn, Estonia.

Liadis, J. S., 2000, GPS TIGER Accuracy Analysis Tools(GTAAT) Evaluation and Test Results, TIGER Operation Branch, Geography Division in the U.S. Census Bureau, http://www.census.gov/geo/www/tiger/gtaat2000.pdf.

Martin, D., 1997, From enumeration districts to output areas: experiments in the automated creation of a census output geography, Population Trend, 88, 36-44.

Martin, D., 1998, 2001 Census output areas from concept to prototype, Population Trends, 94, 19-24.

Martin, D., 2002, Geography for the 2001 census in England and Wales, Population Trend, 108, 7-15.

Marx, R. W., 1986, The TIGER system: automating the geographic structure of the United States census, Government Publications Review, 13, 181-201.

Nael, I. and Nordstrom, L., 1998, Register data, point-based statistics and GIS in the Swedish and the Estonian statistical systems, Working paper No. 26, Work Session in Geographical Information Systems, Conference of European Statisticians, Ottawa, Canada.

O'Grady, K., 2001, A DOQ test project: collecting data to improve TIGER, U. S. Census Bureau, http://www.census.gov/geo/mod/esri_paper.pdf.

O'Grady, K., and Godwin, L., 2000, The Positional Accuracy of MAF/TIGER, U. S. Census Bureau, http://www.census.gov/geo/mod/positional_accuracy.pdf.

Openshaw, S. and Rao, L., 1995, Algorithm for reengineering 1991 Census Geography, Environment and Planning A, 27, 425-226.

Openshaw, S., 1977, A geographical solution to scale and aggregation problem in region-building, partitioning and spatial modeling, Transactions of the Institute of British Geographers, NS 2, 459-472.

Peters, A. and MacDonald, H., 2005, Unlocking the Census with GIS, ESRI Press.

Ratcliffe, J. H., 2001, On the accuracy of TIGER-type geocoded address data in relation to cadastral and census areal units, International Journal of Geographic Information Science, 15, 473-485.

Sobel, J., 1990, Principal components of the Census Bureau's TIGER File, in D. J. Peuquet and D. F. Marble(eds.), Introductory Readings in Geographic Information Systems, Taylor & Francis, 112-119.

Sperling, J., 1995, Development and Maintenance of the TIGER Database: Experiences in Spatial Data Sharing at the U.S. Bureau of the Census, in H. J. Onsrud, and G. Rushton(eds.), Sharing Geographic Information, Rutgers, New Brunswick, 377-396.

Statistics Division, Department of Economic and Social Affairs, United Nations, 2008, Handbook on geographic databases and census mapping(draft), United Nations Expert Group Meeting on Measuring the Economically Active Population in Censuses, April 7-10, New York, USA.

Statistics Finland, 2004, Use of Register and Administrative Data Sources for Statistical Purpose: Best Practices of Statistics Finland.

Tammilehto-Luode, M., Rogstadt, L., Backer, L., 2000, GRID data and area delimitation by definition: Towards a better European territorial system, UN/ECE Work session on Methodological Issues Involving the Integration of Statistics and Geography, Neuchâtel, Switzerland.

Trainor, T., 2003, U.S. Census Bureau geographic support: A response to changing technology and improved data, Cartography and Geographic Information Science, 30(2), 217-223.

U.S. Census Bureau, 2003, 2000 Census of Population and Housing, Summary File 3, Technical Documentation, SF3/11(RV), U.S. Department of Commerce, U.S. Census Bureau, Washington, D.C.

United Nations Economic Commission for Europe, 2007, Register-based Statistics in the Nordic Countries: Review of Best Practice with Focus on Population and Social Statistics.

United Nations, Statistics Division, Department of Economic and Social Affairs, 2004, Integration of GPS, Digital Imagery and GIS with Census Mapping, United Nations Expert Group Meeting to Review Critical Issues Relevant to the Planning of the 2010 Round of Population and Housing Censuses, New York, USA.

Whitman, S., Silva, A., Shah, A., and Ansell, D., 2004, Diversity and disparity: GIS and small-area analysis in six Chicago neighborhoods, Journal of Medical Systems, 28(4), 397-411.

미국 센서스국 웹사이트: http://www.census.gov.
노르웨이 통계청 웹사이트: http://www.ssb.no.
덴마크 통계청 웹사이트: http://www.dst.dk.
스웨덴 통계청 웹사이트: http://www.scb.se.
영국 통계청 웹사이트: http://www.statistics.gov.uk.
유럽연합 통계청 웹사이트: http://ec.europa.eu.
일본 통계국 웹사이트: http://www.stat.go.jp.
핀란드 통계청 웹사이트: http://www.stat.fi.

<부록 1> GIS의 주요 개념

GIS에서 지리 정보(Geographic Information)는 사람, 건물, 산 등과 같은 공간에 존재하는 객체들의 특정한 위치에 대한 정보에서, 나아가 그 위치들 사이의 관계에 대한 정보를 포괄한다. GIS는 이러한 지리 정보를 효율적으로 저장하고 관리하기 위해 하나의 시스템(system)으로 구축되며, 여러 가지 분석 기능을 통해 객관적이고 합리적인 공간의사결정을 지원한다. GIS는 컴퓨터의 개발과 정보 기술의 발전과 더불어, 과거의 종이 지도 제작에서부터, 공간과 관련된 정보 추출에서 분석에 이르는 모든 과정을 자동화한다. GIS는 하드웨어, 소프트웨어, 데이터베이스, 인적자원, 프로시져 등이 복잡한 컴퓨터 네트워을 통해 연결된 구조를 갖는다.

□ 지리참조(geo-referencing): 공간 데이터는 경위도 또는 투영 좌표와 같은 실세계의 수치적 위치 정보를 비롯하여, 우편 구역, 조사구 등 각종 공간 단위들에 대한 지리참조 시스템을 가지고 있다. 이러한 지리참조는 공간 객체의 절대적 또는 상대적 위치뿐 아니라 공간적 관계를 파악하기 위한 기초가 된다.

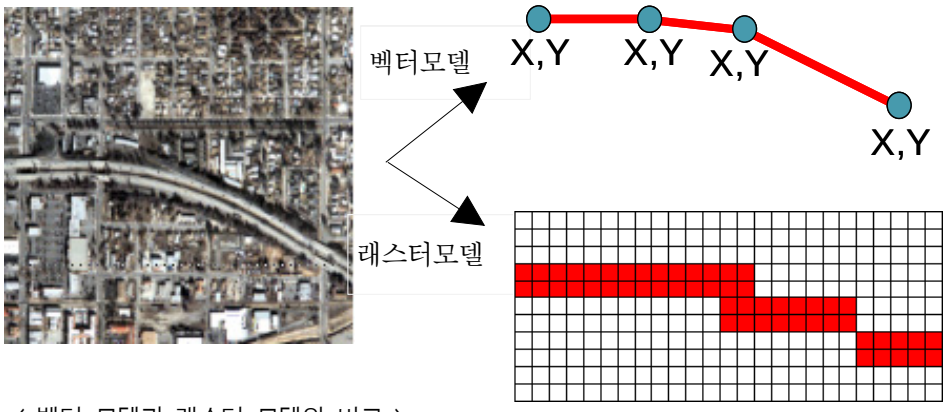
□ 공간적 위상 구조(topology): 공간 데이터는 두 지점의 연결, 인접, 교차 관계 등의 위상 구조를 가짐으로써, 효율적인 공간 정보를 관리하고, 공간 검색 및 분석 능력을 향상시켜 준다.

공간 데이터 모델은 디지털 환경에서 공간 객체와 프로세스를 재현하고 기술하기 위한 방식을 말한다. 오래 전부터 다양한 형태의 데이터 모델이 개발되고 소개되었지만, 이산적 객체(discrete object)와 연속적 필드(continuous field)와 같은 실세계 추상화의 개념적 기반에 근거하여, 일반적으로 벡터(vector)와 래스터(raster) 방식의 논리적 형태를 많이 이용한다.

벡터 데이터 모델은 이산적인, 즉 경계가 분명한 지리적 객체 및 현상을 표현하는데 적합한 모델로, x, y와 같은 실세계 좌표를 이용하여, 점(point), 선(line), 면(polygon)과 같은 기하학적 형태로 실세계 객체들을 재현한다.

래스터 데이터 모델은 지표상의 연속적인 변이를 가진 현상을 재현하는데 적합한 모델로, 실세계를 일련의 셀(cell)들로 분할하고 각 셀들에 위치 및 속성 정보를 할당한다. 각 셀에 위치 정보와 분광 특성을 저장하는 인공위성 이미지가 대표적인 래스터 모델이라 할 수 있다. 아래의 그림은 실세계를 벡터와 래스터 두 가지 상이한 데이터 모델로 재현한 결과를 보여주고 있으며 표는 두 모델의 장단점을 보여주고 있다.

〈 실세계의 벡터와 래스터 재현 〉



〈 벡터 모델과 래스터 모델의 비교 〉

구분	벡터	래스터
데이터 크기	버텍스의 수에 의존	셀크기 의존
데이터 소스	사회, 환경 데이터	원격 탐사, 이미지
응용영역	사회, 경제, 행정	자원, 환경
소프트웨어	벡터 GIS, 자동화된 지도	래스터 GIS, 이미지 프로세싱
해상도	다양함	고정

GIS는 공간 데이터를 효율적으로 관리하고, 유지하는 공간 데이터베이스 기능 외에, 공간 의사결정을 지원하고, 공간 데이터에 부가적인 가치를 부여할 수 있는 각종 분석 기능들을 제공하고 있다.

□ 질의(query)

GIS는 공간 객체들의 위치와 공간적 관계, 그리고 관련된 속성들과의 결합을 통해서 여러 가지 형태의 공간적 질문과 관련된 속성 값의 추출을 위한 질의를 지원한다.

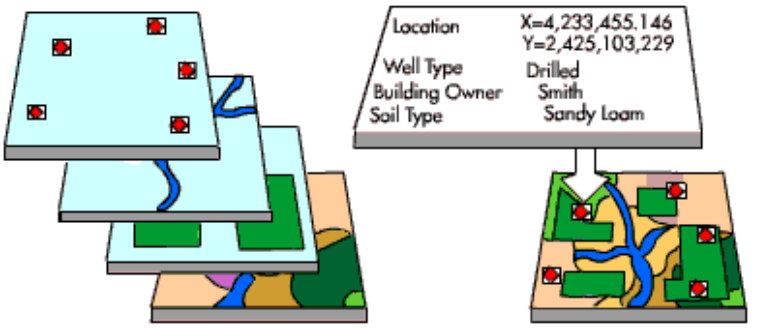
□ 측정(measurement)

GIS는 토지 필지나 특정 구역의 면적이나, 형태, 두 지점 사이의 거리, 도로의 직선거리, 지형의 경사 등의 측정을 손쉽게 계산할 수 있도록 한다. 특히 거리는 많은 공간 분석 응용에 사용되며, 최단직선거리, 네트워상의 경로거리, 구면상의 거리 등 다양한 형태의 측정치들을 계산할 수 있다. 특정 지역의 면적에 대한 계산은 초기 캐나다 GIS의 핵심적 과제였으며, 디지털 환경에서 정확성을 가지고 신속히 계산될 수 있다.

□ 중첩(Overlay)

일반적으로 실세계를 재현하는 GIS 데이터는 동일한 특성을 가지는 하나의 레이어(layer)로 그룹화되어 저장되는데, 각 레이어들은 동일한 좌표체계에서 일치될 수 있기 때문에 중첩을 통해 레이어들 사이의 관련성을 쉽게 분석할 수 있게 된다. 오늘날 GIS의 가장 기본 원리로 인식되고 있으며, 레이어들의 중첩을 통해 합집합, 교집합, 차이와 같은 논리적 연산을 손쉽게 수행할 수 있게 한다.

〈 레이어 중첩 〉

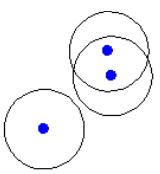


□ 버퍼링(Buffering)

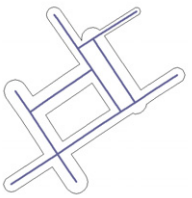
어떤 공간 객체들이 있을 때, 공간적 근접 관계를 확인하기 위해 버퍼를 사용할 수 있다. 특히 버퍼 연산은 원 객체들에서 특정한 거리 내에 있는 모든 지역들을 확인함으로써 새로운 객체들을 생성할 수 있게 한다.

〈 점·선·면 기하학에 따른 버퍼링의 형태 〉

50 m 버퍼



100 m 버퍼, 안쪽 경계 결합하지 않은 상태

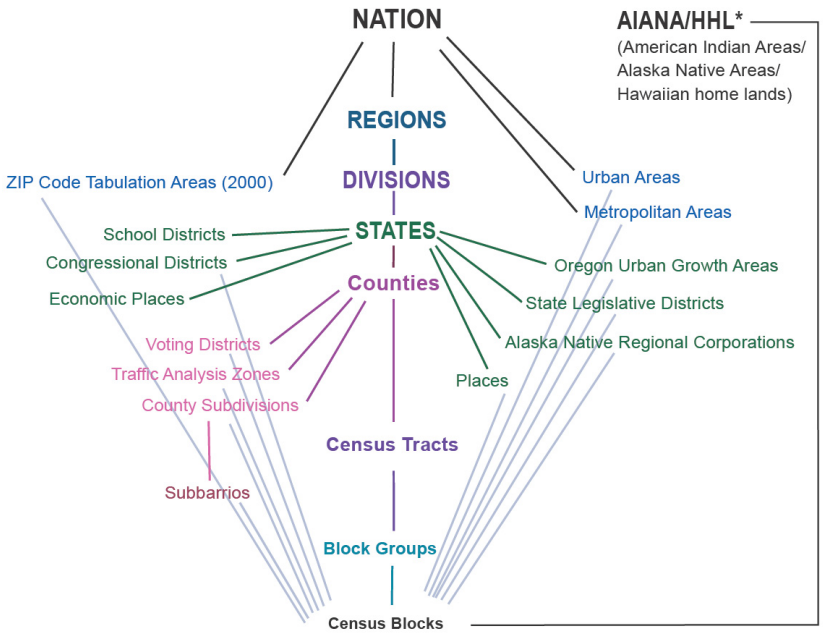


속성 값에 의한 버퍼

<부록 2> 미국 센서스 공간 단위의 표준 체계

센서스와 관련된 공간 객체들은 계층적 구조를 형성하고 있다. 주요한 지리적 계층은 상위 수준에서부터 United States - Regions - Divisions - States - Counties - Census Tracts - Block Groups - Census Blocks 등이다(U.S. Census Bureau, 2003). 미국 센서스의 위상적 공간 객체들에는 연방정보처리 표준 코드(Federal Information Processing Standards: FIPS)라는 체계적인 번호체계가 부여된다. FIPS 코드와 FIPS 코드 문서는 온라인에서 이용가능하다(http://www.itl.nist.gov/fipspubs/index.htm). 이 코드의 목적은 연방 정부의 데이터 자원 사용 능력을 향상시키고, 데이터의 수집·처리·배포에 있어서 불필요한 중복과 불편을 피하고자 하는 것이다. 또한 이 코드를 이용하여 센서스국에서 제공하는 테이블 형식의 센서스 자료와 공간 데이터를 결합할 수 있다.

〈 센서스 공간 객체의 표준 체계 〉



출처: U.S. Census Bureau, 2003.

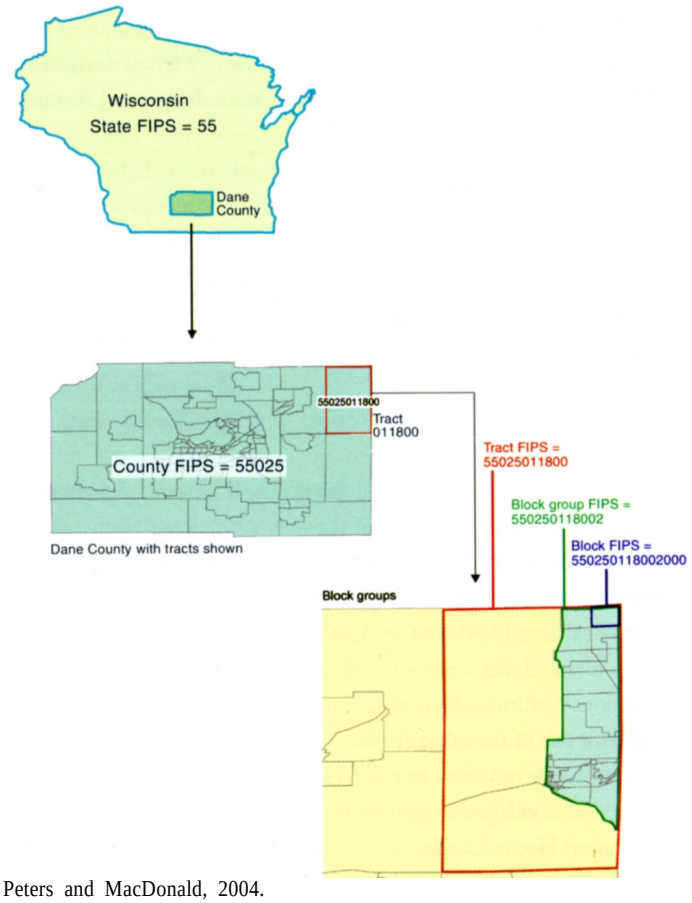
각 주는 알파벳 순서에 따라 두 자리의 FIPS 코드를 부여 받는다. 예를 들어 알바바마(Alabama)는 '01', 아이오와(Iowa)는 '19', 위스콘신(Wisconsin)은 '55'의 FIPS 코드를 갖는다. 각 주에 있는 카운티에는 세 자리 코드가 부여된다. 예를 들어 위스콘신의 데인(Dane) 카운티에는 '025'의 코드가 부여된다. 따라서 각 카운티는 다섯 자리의 코드로 구별되는데, 예를 들어 위스콘신의 데인 카운티는 '55025'이다. 센서스 트랙 코드는 다섯 자리의 주·카운티 코드에 여섯 자리의 트랙 코드가 더해져서 식별된다. 그림에서와 같이 데인 카운티의 가장 북동쪽에 위치한 트랙은 '011800'의 코드를 갖는다. 이 여섯 자리 코드에서 뒤의 두 자리는 시간에 따른 트랙의 분할 내역을 관리하기 위하여 사용된다. 개별 트랙은 블록 그룹으로, 블록 그룹은 다시 블록으로 분할된다. 블록 그룹에는 한 자리의 코드가, 블록에는 세 자리의 코드가 각각 부여된다. 이상에서 설명한 FIPS 코드 부여 체계를 지도를 이용하여 설명하고 있다. 그림에서 보이는 15자리의 FIPS 코드 '550250118002000'에서 각각 55는 주, 025는 카운티, 011800은 트랙, 2는 블록 그룹, 000은 블록을 나타낸다.

어떤 공간 단위는 센서스국의 행정적 필요에서 시작되었으나 데이터 사용자에게 중요해진 경우도 있다. 센서스 블록, BNA(Block Numbering Area) 등은 처음에는 센서스를 수행하고 집계하기 위해 만들어낸 단위였지만 소지역에 통계에 대한 요구가 증대함에 따라 이 단위들은 공식적인 집계와 공표 단위로 널리 사용하게 되었다.

센서스국은 통계적 공간 단위의 이름과 코드를 데이터 공표 시 제공한다. 그러나 그 범위와 위치에 대해서 아는 사람은 거의 없다. 이와 같은 이유로 센서스국은 데이터의 공표와 함께 여러 가지 형태의 지도를 제공한다. 어떤 지도는 특정 공간 단위의 이름과 위치만을 보여준다. 그러나 대부분의 경우는 이들 공간 단위의 경계와 경계 안에 있는 피쳐들의 이름을 포함하기도 한다. 어떤 경계는 강이나 개천과 같은 물리적인 공간을 따라 이루어진다. 통계적 공간 단위는 도로, 철로, 강과 같이 야외에서 쉽게 식별될 수 있는 경계를 필요로 하는데, 이는 원래 조사지역을 할당할 때 조사원이 정확한 범위를 알 수 있게 하기 위해서 사용되었기 때문이다. 이러한 명확하고 쉽게 인식되는 경계는 데이터 사용자로 하여금 통계 단위에 적절한 자료원으로부터 관련 정보를 연계하는 것을 가능하게 해준다.

많은 데이터 사용자들에게 크기(size)는 그 지역의 지리적 범위라기보다는 그 지역 안에 포함되어 있는 사람, 주택, 사업체의 수이다. 그러나 현장 조사의 할당에 있어 지리적 범위는 중요한 고려 요소가 될 수 있으며, 지역의 크기는 데이터 분석에서도 중요한 경우가 있다. 이러한 데이터 분석에 있어서의 요구를 충족시키기 위해 통계국에서는 센서스 데이터의 공표와 함께 그 지역의 면적을 제공하고 있다.

〈 위스콘신 데인 카운티 센서스 블록의 FIPS 코드 부여 사례 〉



출처: Peters and MacDonald, 2004.

형상의 조밀성(compactness of shape)은 통계 단위에서 중요한데, 중심에서부터 경계까지가 대체로 등거리에 있는 것이 바람직한다. 구불구불하거나 찌그러진 지역은 경계 지점들의 통계적 특성이 서로 다를 가능성을 보여준다. 소규모 지역의 데이터는 사용자로 하여금 지리적 단위를 그들의 특정 목적에 맞추어 집계할 수 있는 유연성을 높여주는데, 미국 센서스국의 공간 단위들을 하위 수준에서부터 자세히 살펴보면 다음과 같다.

□ 센서스 블록

센서스 블록은 가로망·도로·하천·철도와 같은 가시적인 지형 형상 및 도시·타운·타운십·카운티의 경계, 가로망이나 도로의 연장선과 같은 가상적인 경계에 의해 정의된 영역이다. 이는 통계 자료의 조사와 공표의 최소 단위로 1940년 센서스에 처음 도입되었으며, 1990년 센서스에서 전국으로 확대되었다. 블록은 타

이거 데이터베이스를 기반으로 인구 규모에 관계없이 구획되기 때문에, 인구가 전혀 없는 블록도 발생한다. 1990년 센서스의 경우 약 700만 개의 블록이 있었는데, 이 중 약 200만 개는 인구가 전혀 없었다(강영옥 외, 2007). 일반적으로 도심 지역에서 센서스 블록의 면적은 작으나 인구 희박지역에서는 수 평방 마일에 이를 수도 있다. 모든 블록에 고유한 ID를 부여하는데, 운영의 효율성 증대를 위하여, 2000년 센서스에서는 집계 블록(tabulation block)과 수집 블록(collection block)에 상이한 번호체계를 사용하였다(1990년에는 단일한 번호체계였다). 수집 블록의 번호는 TIGER/Line에서만 이용가능하며, 데이터 집계에는 사용되지 않는다. 집계 블록은 인디언 보호구역, 센서스 트랙, 선거구, 카운티, 학군, 우편번호 구역 등 어떠한 공간 개체의 경계도 벗어나지 않는다. 집계 블록은 센서스 트랙 내에서 네 자리 번호체계에 의해 식별된다.

□ 블록 그룹

블록 그룹은 센서스 트랙 내에서 네 자리 고유번호 중 첫 자리가 동일한 센서스 블록의 집합으로 하나의 센서스 트랙은 최대 9개의 블록 그룹으로 나뉘어진다. 블록 그룹은 일반적으로 300～3,000명의 인구를 포함하고, 최적의 인구 규모는 1,500명이다. 대부분의 센서스 블록은 센서스국의 통계지역참여 프로그램(Participant Statistical Areas Program)에 참여하는 지역 참여기관에 의해서 결정된다. 지역 참여기관이 없는 지역에 한하여 센서스국이 블록 그룹을 결정한다. 블록 그룹은 인디언 관할기관에 의해서 정의된 것을 제외하고는 주, 카운티 혹은 카운티와 통계적으로 동일한 수준의 경계를 넘어서지 않는다.

□ 센서스 트랙

센서스 트랙은 상대적으로 지속적인 통계 단위로, 센서스국의 통계지역참여 프로그램에 참여하는 지역 참여기관에 의해서 정의된다. 센서스국은 지역 참여기관이 없는 지역에 한하여 트랙을 결정한다. 센서스 트랙의 일차적 목표는 10년 단위 센서스 데이터 표현을 위한 안정적인 지리적 단위를 제공하는 것이다. 2000년 센서스에서부터 미국 전역이 센서스 트랙에 의해 커버되기 시작하였다. 1990년 센서스의 경우, 일부 카운티는 센서스 트랙을 일부 카운티는 BNA로 구분되었다.

센서스 트랙은 일반적으로 1,500～8,000명 사이의 인구 규모를 가지며, 최적 인구 규모는 4,000명이다. 1,500명 이하의 인구를 갖는 카운티 혹은 이와 동일한 수준의 통계 지역도 단일한 센서스 트랙이 될 수 있는데, 인디언 보호구역, 특수 지역(교도소, 군대, 학교, 병원이 포함되어 있는 지역, 노동자 밀집지역 등) 등은 최소

1,000명 이상을 포함해야 한다. 센서스 트랙은 처음 구획될 때 인구 특성, 경제적 지위, 생활 조건에서 상대적으로 등질적이도록 설계된다. 센서스 트랙의 공간적 규모는 거주지 밀도에 따라 달라진다. 센서스 트랙의 경계는 수십 년 동안 유지될 수 있도록 설정되기 때문에 10년 단위 센서스 결과의 통계 비교가 가능하다. 그러나 고속도로, 신도시 개발 등에 의한 도로망 패턴이 물리적으로 변화한 경우 센서스 트랙을 수정하는 것이 필요하다. 또한 센서스 트랙은 인구성장에 의해서 분리되기도 하고 쇠퇴에 의해서 합병되기도 한다.[27]

센서스 트랙은 네 자리 번호에 의해서 식별되며, 두 자리의 하위 번호 체계를 가질 수 있다(예: 6059.02). 소수점 이하의 두 자리는 변화의 이력을 관리하기 위하여 사용된다. 센서스 트랙 번호의 범위는 1~9999까지이고, 카운티 내에서 고유하다.

□ 카운티

카운티는 대부분의 주에서 일차적인 법적 경계인데, 각 카운티 및 통계적으로 동일한 수준의 지역들은 주 내에서 고유한 세 자리의 FIPS(Federal Information Processing Standards) 코드가 부여된다.[28] 이 코드는 대체로 주 안에 있는 카운티의 알파벳 순서대로 부여된다.

□ 주

주는 일차적인 통치상의 경계이며, District of Columbia(DC)는 주와 통계적으로 동일한 수준으로 간주된다. 각 주와 통계적으로 동일한 수준의 공간 개체들은 알파벳 순서에 따라 두 자리의 FIPS 코드가 할당된다. 첫 자리는 각 디비전의 식별자 역할을 한다.

□ 센서스 디비전

센서스 디비전(Census Divisions)은 주와 DC를 그룹화한 것으로, 센서스 지역(Census Regions)의 하위 구역이다. 9개의 센서스 디비전이 있으며, 이들은 센서스 데이터의 표현을 위해 1910년 센서스국이 설정한 것이다. 각 센서스 디비전은 한 자리의 센서스 코드에 의해서 식별된다.

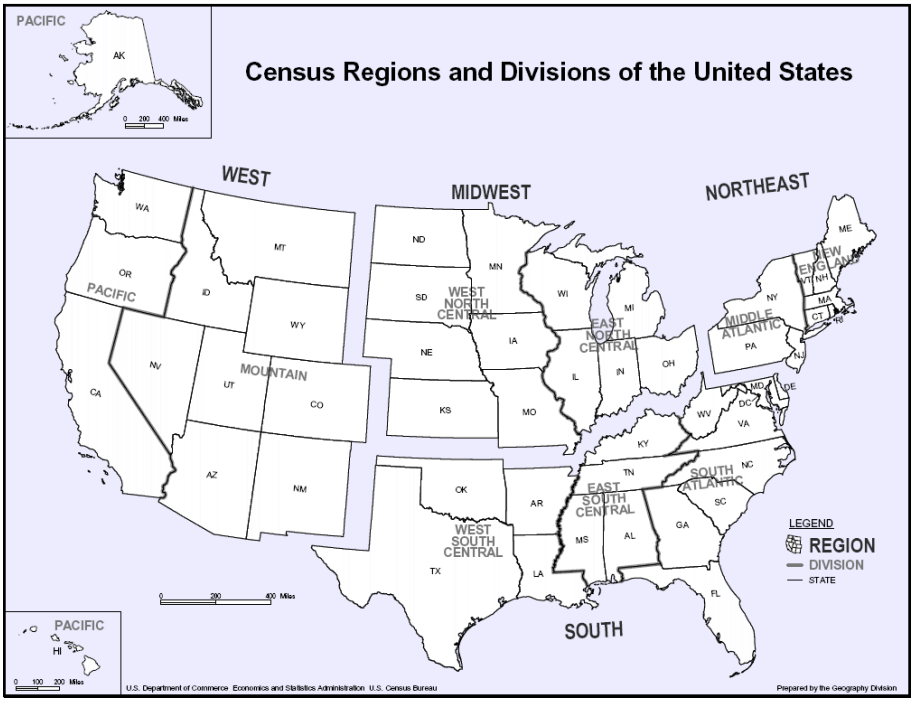
27) 센서스 트랙은 기본적으로 카운티 내에 존재하며, 주와 카운티의 경계는 항상 센서스 트랙의 경계가 된다. 인디언 보호구역의 경우 카운티나 주 경계보다 인구 기준 조건을 우선하기 때문에 예외적으로 2개 이상의 카운티 혹은 주에 걸쳐서 센서스 트랙이 설정되기도 한다. 인구가 최소 기준에 미달하는 센서스 트랙은 다른 센서스 트랙과 병합되는데 이때 기존 센서스 트랙 전체를 다른 센서스 트랙에 병합하는 것을 원칙으로 하여, 기존 구역이 분할되지 않도록 하였다(강영옥 외, 2007).

28) 루이지에나의 parishes/ 알래스카의 census areas, city and boroughs, municipality, organized boroughs/ 메릴랜드, 미주리, 네바다, 버지니아의 independent cities는 카운티와 동일한 수준으로 취급된다.

□ 센서스 지역

센서스 지역(Census Regions)은 센서스 데이터 표현을 위하여 설정된 영역으로, Northeast, Midwest, South, West 네 개가 있다. 네 개의 센서스 지역은 둘 혹은 그 이상의 센서스 디비전을 포함한다. 1910～1940년까지는 North, South, West 세 개의 센서스 지역이 있었고, 1940년 이후 네 개가 되었다. 각 센서스 지역은 한 자리 숫자의 식별자에 의해서 구분된다. 센서스 지역과 디비전의 관계는 다음 그림에 나타나있다.

〈 센서스 디비전과 센서스 지역 〉



출처: www.census.gov/geo/www/us_regdiv.pdf.

<부록 3> 국가GIS 사업

선진국에서는 1980년대 이후 국가 차원에서 공공정보 인프라를 구축하여 정부 업무의 효율화와 국민의 삶 증진을 위한 서비스에 활용하여 왔다. 우리나라에서도 지리정보를 국가의 중요한 사회간접자본으로 인식하여, 1995년, 국가 주도의 NGIS(National GIS) 사업을 추진하였다. 대도시 지역에는 1: 1,000 수치지형도, 그리고 국가 전체적으로는 1: 5,000 수치지형도가 작성되었다(강영옥 외, 2008). NGIS 사업은 국내의 GIS기반을 빠른 시간 내 확충함으로써 이를 통해 국가경쟁력 강화 및 행정 생산성확대에 기반이 되는 새로운 사회간접자본을 구축하는 것이 목적이며, 5년 단위로 현재까지 3차에 걸쳐 사업이 진행 중이다(표 참조).

GIS 관련 업무를 담당하는 기관으로는 국토해양부의 국토정보정책관, 국토지리원, 국립해양조사원, 대한지적공사, 국토연구원, 건설기술연구원, 한국토지공사, 대한주택공사, 통계청, 소방방재청, 산림청, 한국지질자원연구원, 한국도로공사, 한국교통연구원 등이 있다. 현재 공간 정보의 생산은 19개 개별법, 17개 이상의 정부기관에서 개별적으로 생산・관리・활용 중에 있는데, 이러한 분산 관리 문제는 공간 정보 표준화의 장애를 가져왔으며 공간 정보의 공유 및 연계가 부진하고 공간 정보간의 호환성 저하를 가져오고 있다.

〈 국가 GIS 사업계획 개관 〉

구분	1차 NGIS	2차 NGIS	3차 NGIS
기간	1995~2000년	2001~2005년	2006~2010년
비전	경쟁력 강화 및 행정 생산성 제고를 위한 국가공간 정보 구축	국가 공간 정보기반을 확충하여 디지털 국토 실현	유비쿼터스(ubiquitous) 국토 실현을 위한 기반조성
목표 및 추진전략	구각 기본도, 지형도, 지적도 및 국가공간 정보 데이터베이스 구축 공간 정보표준화 및 관련 제도 정비, GIS 소프트웨어 및 DB 툴 개발	국가 공간 정보기반 확충, 국가지리정보 유통망 구축, 핵심 기술 개발, 국민중심 서비스 극대화	GIS 기반의 전자정부 구축, GIS를 통한 삶의 질 향상, GIS를 이용한 새로운 비즈니스 창출, 국가 GIS 활용가치 극대화, 수요자 중심의 국가공간 정보 구축

〈 주요 GIS 관련 연구조직 〉

조직	특징
국토연구원	국토자원의 이용·개발·보전에 관한 정책을 담당하는 국책연구기관 연혁: 국토개발연구원 설립(1978년) 국토연구원으로 개칭(1999년) 국토정보연구센터: GIS 연구 기능 담당
건설기술 연구원	건설기술의 연구 및 정책 개발, 시설물 등에 대한 연구를 목적 유비쿼터스국토연구실: 주요 유비쿼터스 국토 정보 연구 기능 담당
한국토지공사	국토의 효율적인 이용, 개발을 목적으로 설립 국토도시연구원: 주요 연구기능 담당(1처 3연구소 1센터로 구성)
대한주택공사	공공 부문의 주택의 건설, 공급, 관리 목적 연혁: 대한주택공사 설립(1962년) 주택도시연구원: 해당 기관의 연구기능 담당 주택도시정보센터(기술지원부문): 관련 시스템 구축 및 운영
통계청	국가 차원의 통계 생산, 관리, 보급 목적 연혁: 공보처 통계국(1948년), 통계청 승격(1990년) 통계 정보국(통계지리정보과), 통계개발원: GIS 연구 기능 담당
소방방재청	각종 재난으로부터 국민의 생명과 재산을 보호하고, 사회안전망 구축목적 정부조직법과 '재난 및 안전관리기본법' 공표(2004년), 소방방재청 (2004년) 주요 사업: 재해위험 및 홍수관련 자료 등의 DB 구축, 재난관련 종합 정보 제공 및 모니터링 시스템 구현 등
산림청	국토 보전과 국민경제 발전을 도모하기 위한 농림부 산하 기관 주요 사업: GIS를 이용한 산사태 위험지 관련 사업, 국유림영림계획 전산화 프로그램 개발, 산림정보탐사용 항공사진 DB 구축 등
지질자원 연구원	국토와 해외의 지질조사 및 부존자원 연구 목적 연혁: 지질조사소(1918년) 한국지질 조사연구원(2001년) 주요 사업: 지질도 GIS 시스템 구축, 지질자료 표준화 및 유통체계 구축, 해외광물자원 협력 및 기술정보구축, 지상관측 위성자료 유통기구 구축 등
한국도로공사	도로교통의 발달을 위해 설립된 정부투자기관 주요 사업: 고속도로 지리정보형 시스템 DB 구축, 재난도로관리 통합정보 시스템 구축 등
한국교통 연구원	연혁: 교통개발연구원으로 설립(1987년), 개칭(2007) 기능: 교통정책 기술을 연구/개발, 교통 정책/기술에 관련된 국내/외 각종 정보를 수집/조사/분석하여 보급

〈 제2차 국가GIS사업의 주요 담당 조직 현황 〉

구분	담당부처	주요 수행사업
기본지리 정보구축	국토지리정보원	국토공간정보종합관리시스템사업, 기본지리정보구축, 수치지형도 수정, 갱신, 국토모니터링체계구축사업, 국가기준점 정비 및 체계 확립
	행정자치부	지적도면전산화
	국립해양조사원	해안선 조사측량 및 DB구축, 연안해역 해저정보조사, 해양기본도제작사업, 전자해도제작사업, 해양공간 정보시스템구축
	문화재청	문화재 지리정보체계 구축사업
	한국철도공사	철도기본지리정보구축사업
	한국수자원공사	공통유역도 사업
GIS활용 체계구축	통계청	통계지리정보시스템
	농림부	농지정보화사업, 농촌용수 물관리정보화사업
	건설교통부	건축행정정보화, 토지정보망구축사업, 3차원 공간 정보구축 시범사업, 도로와 지하시설공동구축, 국토지반정보DB 구축, 하천지도전산화, 지하수정보관리시스템 구축, 광역상수도종합관리시스템 구축
	광해방지사업단	폐탄광지리정보시스템
	농업과학기술원	농업토양환경정보DB구축사업
	산업자원부	산업단지 수치지도 제작 및 GIS 구축, 산업지리정보체계
	환경부	토지피복분류도구축
	국립환경과학원	자연환경종합GIS-DB구축사업, 화학물질사고대응정보시스템
	산림청	산림지리정보시스템
	해양수산부	항만지하시설물GIS DB 구축, 연안관리정보시스템
	국립해양조사원	연안해양지리정보 실시간제공 시스템사업
지리정보 유통체계 구축	건설교통부	국가지리정보유통체계구축사업
	해양수산부	해양GIS유통체계구축사업
	정보통신부	위성영상정보통합관리사업
국가GIS 기술개발	정보통신부	개방형 LBS컴포넌트 기술개발, 고정밀 위성영상처리 기술개발, 개방형 GIS컴포넌트 기술개발, 3차원 GIS 소프트웨어 개발, 멀티센서공간영상정보통합처리기술개발
국가GIS 표준화	정보통신부	국가GIS 표준화 연구
	국토지리정보원	지리정보표준화사업
	기술표준원	GIS국가표준의 체계확립
	해양수산부	해양GIS표준화체계구축사업
GIS전문 인력양성	건설교통부	국가GIS 인력양성사업
	정보통신부	GIS/ITS 전문인력양성사업
	건설교통부	국가GIS지원연구
	해양수산부	해양GIS정책기반연구

<부록 4> 일본의 격자망 데이터 관련 법령

1973년 7월 12일에 일본 총무성은 통계 목적의 표준 격자망과 격자망 코드에 관한 훈령 143호를 발표하였는데, 이는 다음과 같다.

통계 목적의 표준 격자망과 격자망 코드는 다음과 같다.

1. 표준 격자망

표준 격자망은 기본 격자망(Basic Square), 기본 격자망의 분할을 통해 생성되는 분할된 격자망(Divided Grid Square), 기본 격자망의 통합을 통해 생성되는 통합된 격자망(Integrated Grid Square)으로 구분된다.

(1) 기본 격자망

기본 격자망을 편집하는 방법은 다음과 같다.

일본 지역 전체를 위도 40분 경도 1도로 나눈다

1차 지역 구획을 가로 세로 8등분하여(8x8) 64개의 동일한 크기의 지역으로 나눈다.

2차 지역 구획을 가로 세로 10등분하여(10x10) 100개의 동일한 크기의 지역으로 나누는데, 이는 기본 격자망과 같다.

(2) 분할된 격자망

분할된 격자망은 다음 세 종류의 격자망으로 구성된다. 2분 격자망(Half Grid Square)은 기본 격자망 길이의 2분의 1, 4분 격자망(Quarter Grid Square)은 기본 격자망 길이의 4분의 1, 8분 격자망(Eighth Grid Square)은 기본 격자망 길이의 8분의 1이다.

구분	방법
2분 격자망	기본 격자망을 4개의 지역으로 나눔(2x2)
4분 격자망	기본 격자망을 16개의 지역으로 나눔(4x4)
8분 격자망	기본 격자망을 64개의 지역으로 나눔(8x8)

(3) 통합된 격자망

통합된 격자망은 다음 세 종류의 격자망으로 구성된다.

두 배 격자망(The Double Grid Square)은 기본 격자망의 길이의 두 배이며, 다섯 배 격

자망(The Quintuple Grid Square)는 기본 격자망 길이의 다섯 배, 열 배 격자망(The Decuple Grid Square) 는 기본 격자망 길이의 열배이다.

구분	방법
두 배 격자망	2차 구획 지역을 25개의 지역으로 나눔(5x5)
다섯 배 격자망	2차 구획 지역을 4개의 지역으로 나눔(2x2)
열 배 격자망	2차 구획 지역과 크기가 같음

2. 격자망 코드

표준 격자망 코드는 다음과 같이 정의된다.

(1) 기본 격자망 코드

기본 격자망 코드는 8자리 숫자이다. 처음의 4자리 숫자는 1차 구획을, 다음 2자리 숫자는 2차 구획을, 마지막 2자리 숫자는 3차 구획을 의미한다.

□ 1차 구획은 4자리 숫자에 의해 표시된다. 처음 2자리 숫자는 위도를 의미하는데, 이는 30과 68 사이의 숫자이다. 마지막 2자리 숫자는 경도의 마지막 2자리 숫자를 의미하는데 이는 22와 53사이의 숫자이다.

□ 2차 구획은 2자리 숫자에 의해 표시된다. 하나의 1차 구획 안에는 64(8x8)개의 2차 지역 구획이 있다. 첫 번째 한 자리 숫자는 남단으로부터 북단사이 0에서 시작해서 7로 매겨진 숫자를 의미한다. 마지막 한 자리 숫자는 서단에서 동단사이 0과 7사이의 숫자 중 하나이다. 다시 말해 2차 지역 구획의 코드는 00과 77사이에 있다.

□ 3차 구획은 2자리 숫자에 의해 표시된다. 하나의 2차 구획 안에는 100(10x10)개의 3차 지역 구획이 있다. 첫 번째 한 자리 숫자는 남단으로부터 북단사이 0에서 시작해서 9로 매겨진 숫자를 의미한다. 마지막 한 자리 숫자는 서단에서 동단사이 0과 9사이의 숫자 중 하나이다. 다시 말해 3차 지역 구획의 코드는 00과 99사이에 있다.

(2) 분할된 격자망 코드

분할된 격자망 코드는 다음과 같다.

□ 2분 격자망은 9자리 숫자로 되어 있는데, 이는 8자리의 기본 격자망 코드에 9번째에 한자리 코드를 더한 것이다. 기본 격자망에는 4개(2x2)의 2분 격자망이 있는데, 9번째 자리의 코드는 남서쪽 1 남동쪽 2 북서쪽 3 북동쪽 4의 순서에 따라 숫자가 매겨진다.

□ 4분 격자망은 10자리 숫자로 되어 있는데, 이는 2분 격자망 코드에 10번째에 한 자리 코드를 더한 것이다. 2분 격자망에는 4개(2x2)의 4분 격자망이 있는데, 10번째 자리의 코드는 남서쪽 1 남동쪽 2 북서쪽 3 북동쪽 4의 순서에 따라 숫자가 매겨진다.

□ 8분 격자망은 11자리 숫자로 되어 있는데, 이는 4분 격자망 코드에 11번째에 한 자리 코드를 더한 것이다. 4분 격자망에는 4개(2x2)의 8분 격자망이 있는데, 11번째 자리의 코드는 남서쪽 1 남동쪽 2 북서쪽 3 북동쪽 4의 순서에 따라 숫자가 매겨진다.

(3) 통합된 격자망 코드

통합된 격자망 코드는 다음과 같다.

A. 두 배 격자망코드는 9자리로 기본 격자망의 처음 6자리 코드에 7번째에서 9째 자리 코드가 추가된 것이다. 2차 지역 구획에는 25개 (5 by 5)의 다섯 배 격자망이 있다. 7번째 자리수는 남쪽에서 북쪽으로 가면서 0, 2, 4, 6, 8의 순서로 숫자가 매겨진 값이며, 8번째 자리수는 서쪽에서 동쪽으로 가변서 0, 2, 4, 6, 8의 순서로 숫자가 매겨진 값이다. 9번째 자리의 숫자는 5이다.

B. 다섯 배 격자망코드는 7자리로 기본 격자망의 처음 6자리 코드에 7번째에 한 자리 코드를 추가한 것이다. 2차 지역 구획에는 4개(x 2)의 다섯 배 격자망이 있다. 7번째 자리수는 남서쪽 1 남동쪽 2 북서쪽 3 북동쪽 4의 순서에 따라 매겨진 숫자이다.

C. 열 배 격자망코드는 6자리로 기본 격자망의 처음 6자리 코드와 같다.

<부록 5> 일본 격자망 통계의 종류

- 2005년 인구 센서스의 격자망 통계
(1권: 인구, 가구, 산업)
1. 격자망 지도
    [전국]
    [3대 대도시와 주요 도시] rmf
2.통계표
    [기본 격자망 별 통계표]
    [2분 격자망 별 통계표]
    [4분 격자망 별 통계표(주요 도시)]
3. 제공 데이터
    [도별]
    [1차 지역 구획별]

(2권: 직업, 통근・통학 장소, 이동)
1. 통계표
    [기본 격자망 별 통계표]
    [2분 격자망 별 통계표]
    [4분 격자망 별 통계표(주요 도시)]
2.제공 데이터
    [도별]
    [1차 지역 구획별]

- 2000년 인구 센서스의 격자망 통계
(1권: 인구, 가구, 산업)
1. 디지털 격자망 지도
2. 격자망 지도
    [전국]
    [3대 대도시와 주요 도시]
3. 통계표
    [기본 격자망 별 통계표]
    [2분 격자망 별 통계표]
4. 제공 데이터
    [도별]
    [1차 지역 구획별]

(2권: 직업, 통근・통학 장소, 이동)
1. 통계표
    [기본 격자망 별 통계표]
    [2분 격자망 별 통계표]
2. 제공 데이터
    [도별]
    [1차 지역 구획별]

- 1995년 인구 센서스의 격자망 통계
(1권: 인구, 가구, 산업)
1. 디지털 격자망 지도
2. 격자망 지도
    [전국]
    [3대 대도시와 주요 도시]
3. 통계표
    [기본 격자망 별 통계표]
    [2분 격자망 별 통계표]
4. 제공 데이터
    [도별]
    [1차 지역 구획별]

(2권: 직업, 통근·통학 장소, 이동)
1. 통계표
    [기본 격자망 별 통계표]
    [2분 격자망 별 통계표]
2. 제공 데이터
    [도별]
    [1차 지역 구획별]

- 1990년 인구 센서스의 격자망 통계
(1권: 인구, 가구, 산업)
1. 디지털 격자망 지도
2. 격자망 지도
    [전국]
    [3대 대도시와 주요 도시]
3. 통계표
    [기본 격자망 별 통계표]
    [2분 격자망 별 통계표]
4. 제공 데이터
    [도별]
    [1차 지역 구획별]

(2권: 직업, 통근·통학 장소, 이동)
1. 격자망 지도
    [3대 대도시권]
2. 통계표
    [기본 격자망 별 통계표]
    [2분 격자망 별 통계표]
3. 제공 데이터
    [도별]
    [1차 지역 구획별]

- 2006년 사업체·기업 센서스의 격자망 통계
1. 격자망 지도
   [전국/3대 대도시권]
2. 통계표
   [기본 격자망 별 통계표]
   [2분 격자망 별 통계표]
3. 제공 데이터
   [도별]
   [1차 지역 구획별]

- 2001년 사업체·기업 센서스의 격자망 통계
1. 디지털 격자망 지도
2. 격자망 지도
   [전국/3대 대도시권]
3. 통계표
   [기본 격자망 별 통계표]
   [2분 격자망 별 통계표]
4. 제공 데이터
   [도별]
   [1차 지역 구획별]

- 1996년 사업체·기업 센서스의 격자망 통계
1. 디지털 격자망 지도
2. 격자망 지도
   [전국/3대 대도시권]
3. 통계표
   [기본 격자망 별 통계표]
   [2분 격자망 별 통계표]
4. 제공 데이터
   [도별]
   [1차 지역 구획별]

- 1991년 사업체 센서스의 격자망 통계
1. 디지털 격자망 지도
2. 격자망 지도
   [전국/3대 대도시권]
3. 통계표
   [기본 격자망 별 통계표]
   [2분 격자망 별 통계표]
4. 제공 데이터
   [도별]
   [1차 지역 구획별]

- 2000년 인구 센서스와 2001년 사업체・기업 센서스가 결합된 격자망 통계 1. 격자망 지도 [전국/3대 대도시권] 2. 통계표 [기본 격자망 별 통계표] [2분 격자망 별 통계표] 3. 제공 데이터 [도별] [1차 지역 구획별]
- 1995 인구 센서스와 1996년 사업체・기업 센서스가 결합된 격자망 통계 1. 디지털 격자망 지도 2. 격자망 지도 [전국/3대 대도시권] 3. 통계표 [기본 격자망 별 통계표] [2분 격자망 별 통계표] 4. 제공 데이터 [도별] [1차 지역 구획별]
- 1990 인구 센서스와 1991년 사업체 센서스가 결합된 격자망 통계 1. 디지털 격자망 지도 2. 격자망 지도 [전국/3대 대도시권] 3.통계표 [기본 격자망 별 통계표] [2분 격자망 별 통계표] 4. 제공 데이터 [도별] [1차 지역 구획별]