

GIS 데이터베이스 모델링의 최적화 분석

Optimum of the GIS Database Modeling Method

강인준* · 최현** · 김상석*** · 홍순현****

Kang, In Joon · Choi, Hyun · Kim, Sang Suk · Kang, Ho Yoon

1. 서론

1.1 연구동향

과거의 GIS는 연구도구의 차원에서 활용되었으며 데이터베이스 설계가 정보의 통합적인 분석, 관리, 계획, 의사결정 및 정보의 공유 목적에 맞게 설계되기보다는 부분적인 주제에 맞게 설계되어 왔다. 현재는 도시행정 및 국가기관과 같은 조직의 정보를 통합하여 의사결정과정을 지원하는 시스템 통합 도구로서 그 발전방향이 바뀌고 있다. 정부는 GIS 데이터베이스 구축의 필요성을 인식하고 1995년부터 정보화 시대의 새로운 사회간접자본으로서 공간정보를 정착시키기 위한 노력을 시작했는데, 그 첫 단계로 '국가지리정보체계 구축 기본계획'을 수립하였다. 이를 토대로 공간정보에 대한 인식확산, 공간정보의 보급, 기초 기술개발을 위한 여러 가지 연구가 수행되었다. 그 중의 하나로서 국립지리원의 수치지도 제작은 공식화된 공간 정보의 보급이라는 점에서 큰 의미가 있다(수치지도 데이터 모델 연구, 국립지리원). '95년도부터 시작된 국가GIS구축사업(NGIS)과 정부의 정보화공공근로사업에 힘입어 GIS 기반의 다양한 형태의 주제도가 제작되고 있는 실정이다. 이러한 정보화의 추진에 힘입어 수자원분야에서도 다양한 형태의 하천정보의 구축을 위한 하천 GIS를 비롯한 다양한 주제도 구축 사업이 추진되고 있다. 이러한 하천 정보 관련 다양한 GIS 주제도의 원활한 구축과 활용을 위하여 비용 효과적인 주제도의 구축방법의 정립과 함께 구축된 하천정보를 보다 용이하게 많은 사람이 공유하여 활용할 수 있도록 제반 표준화 관련 사항이 정립되어야 한다. 이러한 표준화는 세부적으로 표준적인 하천정보의 분석 절차를 위한 표준분석 방법론, 하천정보의 용이한 추출과 효율적 관리를 위한 표준모델, 자료의 원활한 공유를 위한 자료의 표준 포맷, 자료의 효율적 검색과 공급을 위한 표준 메타데이터의 설계, 그리고 하천정보의 원활한 공급을 위한 사용자 요구분석에 기반 한 향후 GIS 기반의 하천정보의 공급 방안이 제시되어야 한다. 이러한 표준화 관련 전반적인 사항의 정립을 위하여 '하천정보 표준화에 관한 연구(한국수자원공사)'가 제작되었다.

1.2 연구 목적

현재 국내에는 본 연구와 유사한 분야에 대한 경험을 가지고 있는 전문가가 얼마 되지 않고, 또 비슷한 문헌도 부족한 형편이다. 부산시 UIS(과업 기간 : 1999년 12월 30일 ~ 2002년 6월 29일)의 경우 우선 도시기반 시설물에 대한 통합적이고 합리적인 자료에서 쓰기 위해서, 그리고 업무 프로세스 개선을 통한 업무 효율화와 마지막으로 유관업무 정보의 통합관리를 통한 행정의 과학화를 위해서 추진되었다. 따라서 본 연구는 국내의 기존 GIS 데이터베이스 모델링 기법들이 가진 장·단점을 파악하여 문제점을 해결함으로써, 활용의 효율성을 최대한 높이는 데 있다. 그리고 도시시설물의 근간이 되고 행정기관 뿐 아니라 도시가스, 통신, 전기 등 타 유관기관 시설물과 밀접하게 관련되는 도로시설물 데이터베이스에 대한 모델링을 통해 합리적이고 효율적인 도로기반시설물 데이터베이스와 이것을 공유하기 위한 도로 메타데이터를 정립하고자한다.

1.3 연구 범위 및 방법

- * 정회원 · 부산대학교 공과대학 토목공학과 교수 · 공학박사 · 051-510-2353(E-mail: ijkang@pusan.ac.kr)
- ** 부산대학교 생산기술 연구소 연구원 · 051-510-3209(E-mail: xhyun@pusan.ac.kr)
- *** 부산대학교 토목공학과 박사과정 · 051-510-2188(E-mail: civil-1614@hanmail.net)
- **** 밀양대학교 이공학부 토목공학과 조교수 · 055-350-5316(E-mail: hongsh@arang.miryang.ac.kr)

본 연구는 우선, 국내 지형도 모델 분석을 하였다. 즉 기존에 국립지리원에서 추진되고 있는 국가지형도 표준모델의 추진 상황을 파악하고 모델의 특성을 분석 정리하였다. 아울러 개념적, 논리적, 의미적 모델의 구성 범위를 파악하고, 모델의 구성 시 전제된 제약사항, 기본 가정 및 요구사항을 파악하였다. 이론적 데이터 모델의 구성에 있어서 특히 데이터 셋과 도엽 및 Theme의 체계, 지형지물의 체계를 포함하여 분석하고, 국내의 모델에서 정립된 코드 체계 및 feature set 등을 분석하였다. 이와 함께 국내의 모델에서 제시하는 위치 원시요소와 관계형 원시요소를 파악하였다. 그리고 국립지리원의 수치지도 작성 작업내규와 대한측량협회의 지도도식 규칙을 참고 하였다. 지리정보 데이터베이스 설계 지침과 한국수자원공사를 비롯한 국가지리정보체계 구축사업의 결과물로 나온 연구보고서들을 검토하였으며, 마지막으로 부산시 UIS 도로분야 데이터베이스 분석을 통해서 도로 데이터베이스 모델링과 메타데이터 정립을 수행하였다.

2. 본론

2.1 GIS 데이터베이스

2.1.1 데이터베이스 정의

데이터베이스는 다양한 정보를 적시에, 적절하게 공급함으로써 정보요구를 충족시키는 틀로써, 어떤 조직체에서 모든 응용프로그램이나 사용자의 편의성을 위해 통합되고 저장됨으로써 언제든지 운영 가능하게 공유할 수 있는 데이터의 집합체이다. 이러한 데이터베이스의 데이터는 체계적이고 규칙적으로 구성되어 조직의 업무처리에 필요한 정보를 체계적으로 제공한다.

2.1.2 데이터베이스의 특성

데이터베이스의 데이터들은 중복을 최소화하여 기존 파일 처리 방식체계에서 나타나는 문제점을 미리 방지하고 내정된 사용자에게 접근할 수 있는 권한을 부여함으로써 데이터 보안이 가능하다. 또한 실시간 내에 접근이 가능하고 계속적인 변화 속에서도 정확성을 유지할 수 있으며 여러 사용자들이 함께 공유할 수 있는 특성을 가진다.

2.1.3 데이터 모델링

데이터베이스 설계에 있어서 가장 중요한 것 중 하나는 사용자의 사용 관점을 실세계 모델링하는 부분에서 충실하고 완전하게 어떻게 반영하느냐 하는 것이다.

2.1.3.1 데이터 모델링 개념

데이터 모델링의 개념은 논리적 데이터 모델링과 물리적 데이터 모델링으로 나눌 수 있는데 논리적 데이터 모델링의 경우 업무의 의사소통 도구인 ERD(Entity Relation Diagram)와 시스템의 구성과 무관하게 진행과 M:M관계, Arc관계, Recursive관계, Super/SubType등이 많은 것을 들 수 있겠고, 물리적 데이터 모델링의 경우 성능관점에서 논리적 모델을 재구성 · 시스템 환경 및 애플리케이션의 구조를 고려 · 비정규화(Denormalize) 작업 필요 · 데이터베이스의 오브젝트 활용 등을 설명할 수 있어야 한다.

2.1.3.2 데이터 모델링 절차

데이터 모델링 절차는 크게 표 1과 같은 절차를 따라야 한다.

표 1. 데이터 모델링 절차

단계	내용
1	엔터티(Entity) 도출
2	엔터티 관계(Relationship) 설정
3	유일키(Unique Identifier) 설정
4	속성(Attribute) 설정
5	정규화(Normalization)
6	비정규화(Denormalization)

2.2 수치지도

국립지리원 수치지도는, 1990년 중반부터 수치지도 제작을 위한 중복투자를 방지하기 위해 제작되었다. 현재 국립지리원 수치지도는 다양한 분야에서 기본도로 사용되고 있는데, 특히 지도제작, GIS 데이터베이스, 주제도 제작 등에 사용된다. 현재 지자체에서 도시종합시스템 기본도로 사용하고 있는 1:1,000 수치지도의 분류체계는 크게 세 가지로 나누어지는데, 7개 그룹의 대분류, 30그룹의 중 분류, 239그룹의 소 분류로 나누어진다. 7개 그룹의 대분류에는 시설물, 수계, 지형/지질, 식생, 행정/경계, 지적, 일반이 있다. 위 분류체계의 특징은 지형지물을 그룹화 하는 기준이 "지형지물의 기본성격" 중심이며 중복 정의된 지형지물, 정의수준이 일정하지 않은 것 및 속성으로 표현해야 할 것이 있다.

2.3 프레임워크 데이터

프레임워크 데이터는 "응용분야에 따라, 여러 종류의 지리정보를 구축하거나 활용할 때, 공통적으로 참조되거나 이용되는 기본적인 데이터와 체계"를 말한다. 예를 들어, 상하수도를 위한 지하매설물 데이터베이스를 구축할 때나 통신시설물 데이터베이스를 구축할 때, 도로에 관한 데이터는 지하매설물의 상대적인 위치를 결정하기 위해 공통적으로 참조되는 데이터를 프레임워크 데이터라고 할 수 있다.

프레임워크 데이터는 모든 지리정보시스템 데이터베이스의 중심이 되는 부분이며 경제적 데이터베이스 구축, 여러 주제도사이의 불일치 제거, 여러 데이터베이스의 통합 연결고리의 세 가지 기능을 제공한다. 프레임워크 데이터는 지형지물 유일 식별자(Unique Feature Identification)와 핵심 지형지물(Core Feature Set), 마지막으로 프레임워크 데이터 모델과 같은 내용을 표현하는 구성요소를 가지고 있어야 한다. 표2는 외국 프레임워크 데이터의 지형지물 목록(표 2)

표 2. 외국 프레임워크 데이터의 지형지물 목록

수치지형도의 대분류	미국	영국	호주
철 도	-	-	철도중심선
하 천	수문	-	해안선, 하천, 내륙수계
도 로	교통	도로중심선	도로중심선
건 물	-	-	주요시설물, 항공시설, 해양교통시설
지 류	-	-	-
시 설 물	-	-	-
지 형	측지기준 표고데이터	측지기준 표고	측지기준 지표면 표고
행정경계	행정구역	경계선	행정경계
주 기	-	-	지형지물명/지명
기 타 (외국에만 존재)	수치정사영상, 지적	지형도, 좌표화 된 주소 정보	토지이용현황도, 항공사 진, 인공위성데이터

2.4 메타데이터

메타데이터(Metadata)는 "데이터에 대한 데이터"와 데이터에 대한 이력서 및 데이터 유통 시 데이터에 대한 이해를 높이고 정보의 활용을 촉진하는 중요한 기능을 담당하고 있어야 한다. 메타데이터는 수치지도 공급 체제 구축을 위한 기반을 마련하고 수치지도 분배 업무의 효율화와 신속한 자료 검색 환경 조성뿐만 아니라 타 시스템과 연결할 수 있는 설계 기반을 마련할 수 있어야 한다. 각 지자체의 메타데이터 관리 시스템의 구축 방향은 향후 각 지자체의 지리정보관련 부서가 FGDC의 Clearing house와 같은 기능을 수행할 수 있도록 그 기반을 마련하는 것이다. 즉, 각 지자체 내의 지리적 영역을 갖고 추진하고 있거나 구축되어 있는 시스템(예: 하수도, 상수도, 도시계획, 지적 등)의 데이터베이스를 통합관리 하고 유통할 수 있는 체제를 만드는 역할을 한다. 메타데이터 표준은 NGIS 표준화 분과인 한국 전산원에서 관장하고 있으며 '국가지리정보체계(NGIS) 정보유통을 위한 정보기록방식(메타데이터) 표준화를 위한 연구' 결과 우리나라는 ISO/TC 211의 메타데이터를 채택하여 이용하는 것이 여러 면에서 타당하다고 판단하였으며, 이러한 결정에는 미국 연방 정부 (FGDC)의 메타데이터를 작성한 Mr. Recharad Hogan의 조언도 고려되었고 미국도 앞으로는 ISO/TC 211의 메타데이터에 적합하도록 메타데이터를 수정할 방침이라고 하였다. 우리나라의 경우 각 기관에서 메타데이터 연구와 실제 구축 예가 있으며, 그 바탕은 크게 FGDC표준과 ISO/TC 211 표준에 근거를 두고 있다. 메타데이터 표준 설정에 대한 대표적인 국내적용 사례와 전개는 그림 1과 같다.



그림 1. 국내적용 사례와 전개

2.5 국립지리원 수치지도 데이터 모델 분석

수치지도 데이터 모델에 관한 연구에서 제시한 모델은, 국립지리원에서 제작하고 있는 수치지도를 저장, 관리하기 위한 모델(NGI 데이터 모델)이다. NGI 데이터 모델에서는 실세계의 주요 지형지물에 대하여 단일 식별자(UFID)를 부여한다. 또한 데이터의 내면적, 명시적 위상 정보를 생성함으로써, 기존 그래픽 형태의 데이터보다 데이터의 활용성을 높이려 한다. NGI 데이터 모델에서 데이터의 구성 및 저장단위는 도엽이며, 각 도엽은 한 개 이상의 주제(레이어)로 구성된다. NGI 데이터모델에서 정의하고 있는 데이터의 형태는 벡터 데이터와 격자 데이터, 영상데이터가 있다. 그림 2는 수치지도 데이터 모델의 개념적 모델은 도엽과 주제, 사상 간의 관계를 보여준다.

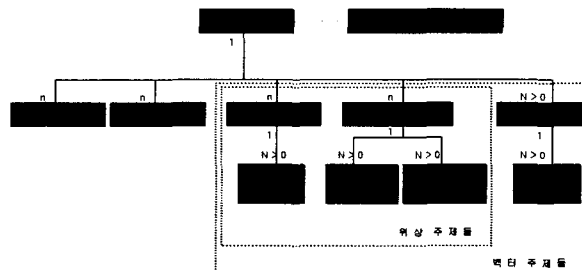


그림 2. 수치지도 데이터 모델의 개념적 모델

국립지리원 수치지도 데이터 모델에서 정의하고 있는 공간객체의 유형에는 실세계의 현상을 표현하는 사상, 사상의 기하적 구성요소인 기하 공간 객체, 공간좌표로서 위치를 나타내는 위치 객체, 지형지물의 위상적 관계를 명시적으로 나타내는 관계형 공간객체(위상 공간객체)가 있다.

2.5.2 수치지도 데이터 모델에 관한 연구

그림 3은 단순 사상 모델은 단순사상이 기하 공간 객체인 폴리라인과 점에 의하여 표현됨을 보여준다. 단순사상을 구성하는 폴리라인은 교차하지 않으므로 선형지형물이 다른 지형물을 가로질러 갈 수 없다.

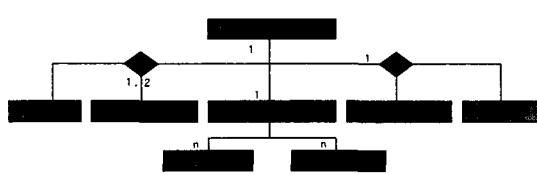


그림 3. 단순 사상 모델

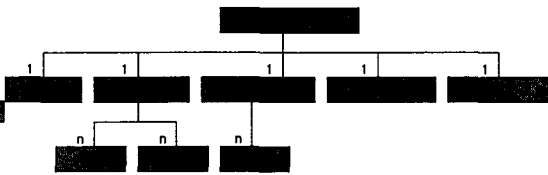


그림 4. 네트워크 세그먼트 사상 모델

그림 4는 네트워크 세그먼트 사상모델은 도로나 전력, 송전선, 강, 철로와 같은 선형사상의 움직임이나 흐름을 표현하는데 사용된다. 그림 5는 교차 사상모델은 선형 사상이 교차하는 지점을 표현해준다. 교차 사상은 특정 교차점에 대하여 연결 매트릭스를 가질 수 있다. 교차점은 교차점에서 만나는 세그먼트의 목록에 대한 정보를 가지고 있다.

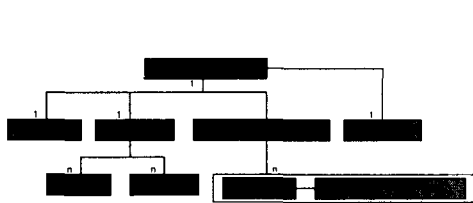


그림 5. 교차 사상 모델

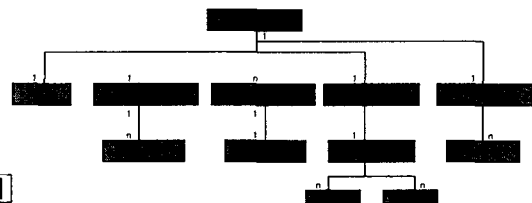


그림 6. 면 사상 모델

그림 6은 면 사상 모델에서 면은 각각의 폴리곤이 독립적으로 분리되어 있고, 회형적 위상(explicit topology)이 없는 이중입력(duplicated geometry)과, 외형적 위상을 갖도록 경계선 또는 공유경계선의 관계를 만족하는 단일입력(single geometry)으로 정의된다. 그림 7은 기하 공간 객체 모델은 기하 공간 객체들 간의 관계를 보여준다. 기하 공간 객체의 종류에는 점, 폴리라인, 폴리곤이 있다. 폴리곤은 기하 공간 객체의 형태로도 존재하지만 관계형 객체로도 존재한다. 그림 8은 관계형 사상을 모은 위상 공간 객체들간의 관계를 보여준다. 관계형 공간 객체는 위상주제에서 지형지물들을 명시적으로 상호간에 연결시켜야 하는 경우에 사용된다. 관계형 공간객체는 지형지물 코드나 속성 등을 갖지 않는다.

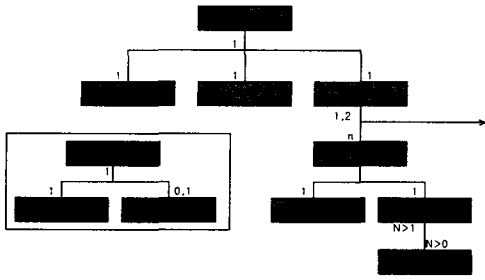


그림 7. 기하 공간 객체 모델
(폴리라인과 좌표, 점의 관계)

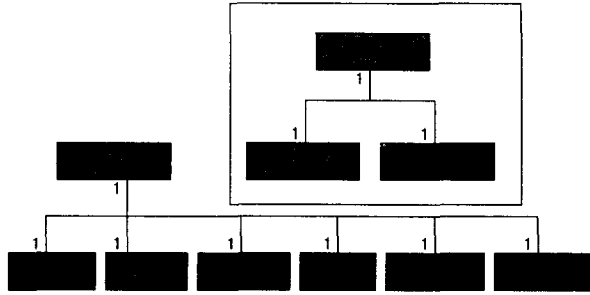


그림 8. 관계형 사상 모델(링크와 점의 관계)

2.6 지리정보 데이터베이스 설계 표준지침 분석

한국정보통신기술협회에서 발간한 '지리정보 데이터베이스 설계 표준지침 분석'은 서로 다른 지리 정보 커뮤니티 간에 지리 정보의 상호운용성을 확보하고, 기관이나 지자체 등의 GIS 수요기관에서 지리 정보 DB를 구축 시 활용할 수 있는 지리 정보 스키마 설계 지침을 제공하고 있다. 이 표준은 지리 정보 DB의 스키마를 설계하기 위해서 표 4와 같이 크게 두 개로 나눈다.

표 3. 표준의 구성 및 범위

단계	내용
응용 스키마 설계 지침	<ul style="list-style-type: none"> ○ 응용 스키마 설계 규칙은 ISO/TC211의 19109를 기반으로 함 ○ GIS 수요기관의 공통 기본도 DB의 개념적 모델 설계를 위한 지침 ○ 개념적 스키마 언어(예를 들면 UML)를 이용 ○ 생성된 응용 스키마는 ARC/INFO, Gothic, MGE 등에 독립적인 형태의 추상 명세를 제공 ○ 예를 들면, GIS 응용과 GIS S/W에 독립적인 지리 정보 DB를 위한 ER(Entity-Relationship) 스키마를 만드는 것과 동일함 ○ 본 표준은 공간 데이터 포맷을 제공하는 것이 아님 ○ 서로 다른 GIS 수요기관의 지리 정보 DB로 쉽게 변환되도록 하기 위한 공통 지리 정보 DB 스키마 설계 표준을 제공함
구현 스키마 설계 지침	<ul style="list-style-type: none"> ○ 응용 스키마 생성 규칙에 따라 생성된 응용 스키마로부터 각 지리 정보 DB 스키마로 변환하는 절차를 제공 ○ 구현 스키마의 설계 규칙의 내용 <ul style="list-style-type: none"> ▷ 일반 피쳐 모델 변환 ▷ 공간 데이터 모델 변환 ▷ 일반 피쳐 모델의 매핑 ▷ 공간 원시 요소의 매핑 ▷ 원시 요소 매핑

응용 스키마는 두 가지 목적을 담당한다. 첫 번째는 특정 응용 분야에서 데이터 구조를 공통되고 정확하게 이해하는 것이다. 두 번째는 자동화된 메커니즘을 적용하여 데이터 관리를 하기 위한 컴퓨터가 해독할 수 있는 스키마이다.

2.7 비교고찰

GIS(Geographic Information System)에서 데이터베이스는 일반적인 업무시스템 및 분석시스템의 질을 결정하는 매우 중요한 요인이다. 또한 많은 시행착오와 노력이 필요하며 구축하고자 하는 시스템의 구조(Architecture, H/W, S/W)의 지원여부 및 데이터 모델링 대상인 실제 세계의 여건에 따라 많은 제약사항이 발생하여 논리적으로 원하는 업무기능 및 분석을 수행하는 데 장애 요인이 될 수 있다. 따라서 GIS 데이터베이스의 설계는 기존의 문자정보만을 활용하여 정보화한 시스템에 비해 보다 복잡적이고 신중하게 계획되고 수행되어야 한다. 본 연구를 통하여, 일반적으로 데이터모델의 검증을 위해서는 모델의 구현과 수정작업이 반복적으로 이루어져야 하는데, 모델에 대한 프로토타입의 구현 뿐만 아니라, 모델에서 표현할 수 없는 특수한 지역이 현실적으로 존재하지 않는지에 대한 검토가 필요하다. 그리고 연구를 수행하면서 많은 예외적 상황을 가정하고, 모델의 표현력을 검토하였지만, 현실적 제약으로 인하여 모든 지역을 현지 조사할 수 없는 어려움이 있었다. 마지막으로 현재 대부분의 데이터모델링 방향이 역공학(Re-Engineering)방식으로 모델링이 이루어지고 있었다. 이는 기존에 구축되어 운영중인 시스템의 데이터베이스와 호환성을 유지하기 위해서 현실을 이론적으로 잘 반영하기 위한 이론적 근거가 미약하다는 것을 알 수 있었다.

3. 결론

본 연구는 도로시설물 데이터베이스에 대한 모델링을 통해 합리적이고 효율적인 도로기반시설물 데이터베이스와 이것을 공유하기 위한 도로 메타데이터 정립에 있으며, 국립지리원 수치지도 데이터 모델 · 지리정보 데이터베이스 설계 표준지침(한국정보통신기술협회) 및 하천정보 표준화에 관한 연구(한국수자원공사)를 분석하였다. 본 연구를 수행한 결과 다음과 같은 결론을 얻었다.

첫째, 각 연구기관이나 정부기관에서는 데이터모델링으로서 객체지향형 데이터모델을 사용하였는데, 분산환경과 개방형 시스템을 구축하는데 효과적인 뿐만 아니라 실세계 현상에 대한 표현력이 우수하다는 것을 알 수 있었다.

둘째, 도로시설물 데이터베이스에 대한 모델링 과정에서 관련 연구 및 국내외 자료들의 분석결과, 설문조사, 현 수치지도 분석 등을 수행하고, 작업의 검증을 위해서 국내외 자문위원들로부터 조언을 구하여야 한다.

셋째, 모델링 기법에서 모델의 요구사항은 모델의 데이터 표현력과 데이터 관리 및 서비스측면에서 정의하여야 하는데, 데이터 표현력을 위한 요구사항에는 데이터는 Feature 단위로 정의하여야 할 것이다. 그리고 도로시설물을 표현하는 데이터, 도로데이터의 위치기준 데이터, 도로관련 업무에서 공통적으로 사용하는 데이터들을 조사하여 '표준 도로 핵심데이터'를 정의하여야 한다.

넷째, 연구 결과 검토를 위해서는 논리 모델에 대한 물리모델을 작성하여 실제 시범구축을 할 필요가 있는데, 시범 구축된 데이터를 실제 도로 GIS분야에 응용하여 이 데이터로부터 사용자 GIS 구축시 보완해야할 사항을 검토해야 한다. 그리고 기본 프레임워크 데이터를 위한 핵심지형지물을 연구하였는데, 여기에 속하는 지형지물들을 어떻게 도면에 표현할 것인지에 대한 지침이 정해져야 할 것이다.

참고문헌

1. 한국수자원공사, "하천정보 표준화에 관한 연구", 2001. 6
2. 건설교통부 국립지리원, "수치지도 Data Model 연구(II)", 1999. 11
3. 건설교통부 국립지리원, "수치지도 작성작업규칙", 1995. 5
4. 서울 시정개발 연구원, "서울시 GIS 기본도 구축을 위한 기술 지침 연구", 1996
5. 한국전산원, 국가 기간 전산망 표준화 연구 중 객체지향 데이터베이스 표준화에 관한 연구", 1994.12
6. 국토개발연구원, "국가 기본도 수치지도화 방안 연구", 1996. 5
7. ESRI, "Modeling our World", 1999. 11

9. Department of Computer Science Stanford University, 'A First Course in Database Systems', 1997
10. ODMG-93, 'The Object database standard', 1994
11. Hye-Kyong Kang, Soon-Hee Do, and Ki-Joune Li, 'Model-Oriented Generalizat Rules', Proc. ESRI Conf. 2000.
12. A.Y.Tang, T.M.Adams, E.L. Usery, 'A spatial data model design for feature-based geographical information systems', Journal. IJGIS, Vol.10, No.5, pp.643-659, 1996
13. M.Sester, K.H.Anders, V.walter, 'Linking Objects of different spatial data sets by Integration and aggregation', Journal GeoInformatica, Vol. 2, No. 4, pp.335-358, 1998