

GML(Geography Markup Language) 응용: 지질주제정보 데이터 모델개발과 시험구현

이기원, 문선희*, 지광훈

한성대학교 정보공학부, 서울대학교 지구과학교육과 대학원,
한국지질자원연구원 국가지질정보센터

GML Application: Developing of Data Model for Geological Information and Its Implementation

Kiwon Lee, Sun-Hee Moon, Kwang-Hoon Chi

요약

OGC(Open GIS Consortium, Inc.)에서 2000년에 XML의 지리정보 처리를 위한 GML 1.0 이 제안된 이후 최근 GML 3.0이 발표되기까지 국제적으로는 다양한 GIS 응용분야에서 GML을 현실 문제에 이용하고자 하는 시도가 이루어지고 있다. 현재 우리나라에서도 국가 GIS(NGIS)에서 GML을 도입하여 국내 표준안으로 추진하고자 다양한 기반연구가 진행되고 있다. 본 연구는 중요한 GIS 응용중의 하나인 지질자원 정보 분야에서 GML을 활용하기 위한 기반연구로서 우리나라의 표준 지질도(또는 수치지질정보)와 기타 지질자원 정보 인프라 구축에 적용될 수 있는 다양한 지구과학자료를 대상으로 하였다. 따라서 본 연구에서는 지질자원 분야에서의 GML 응용을 위한 기반 데이터 모델과 이를 처리하기 위한 아키텍처를 UML 기법으로 설명하고, 이를 바탕으로 시험 구현된 운용환경에서 실제 대상 지역을 선정하여, 연구지역에 대한 지질자원 정보의 GML 적용 사례를 제시하고자 한다. 이를 위한 연구 내용으로는 수치 지질도를 XML-GML의 기본 스키마구조에 따라 인코딩한 뒤, 주제정보의 표현을 위한 XSLT, SVG 파일 변환 처리 등을 수행하였다. 결론적으로 본 연구는 지질자원 분야의 정보화 인프라 구축을 목적으로 하는 기초 시험연구로 수행되었는 바, 우리나라에서는 2001년 이후 GIS 수치 주제도로써 수치 지질도의 보급이 활성화되면서, 이 정보를 제작/공급하는 기관이나 업무에 사용하는 기관 등에서 GML 기반으로 데이터를 제작하거나 운용이 가능할 것으로 생각된다. 또한, 다른 분야에서 GIS 주제정보를 다루는 경우에도 GML을 주제정보와 연계하여 현실 문제에 적용이 가능한 참조 모델로 이용될 수 있을 것으로 생각된다.

1. 서론

현재 GIS 데이터 공유에 대한 국제적인 동향은 웹 상에서 공유가 가능한 포맷에 대한 시도가 활발하게 이루어지고 있으며, 이를 위해 OGC(Open GIS Consortium)에서는 웹 상에서 광범위하게 사용되고 있는 XML(Extensible Markup Language)을 활용한 GML(Geography Markup Language)을

지리 정보의 웹 표준으로 제안하였다.

GML은 지형지물의 공간적 특성과 비공간적 특성을 포함한 지리정보의 저장 및 교환을 위한 XML 인코딩으로 인터넷 환경에서의 공간정보의 저장과 전송을 가능하게 하고, 공간 데이터를 기반으로 하는 다양한 분석 작업에 맞게 확장할 수 있게 한다. 또한 전송의 효율을 위한 압축이 가능하게 하고 데이터 표현에 관하여 공간정보와 비공

간정보의 구분이 가능하도록 한다는 기본적인 설계 목적을 가지고 있다. 2000년에 GML 1.0이 제안된 이후 최근 GML 3.0이 발표되기까지 국제적으로 다양한 GIS 응용 분야에서 GML을 현실 문제에 이용하고자 하는 시도가 이루어지고 있다. 현재 국내에서도 기본지리정보 구축사업, 국가지리정보 유통메타데이터 등의 공동활용을 위한 데이터 교환 포맷으로 GML이 제안된 상태이며 국내 표준안으로 추진하고자 다양한 연구가 진행되고 있다.

GML은 국제표준으로서 데이터 모델의 구조(Schema or Syntax)와 의미(semantics)를 지원하며, XML(eXtensible Markup Language)을 기반으로 하고 있으므로 데이터 포맷에 대한 전문가가 없어도, XML 경험이 있는 프로그래머들에 의해 데이터 포맷이 이해될 수 있다. 또한 원래 데이터의 구조를 그대로 유지한 채, 이와는 별개로 사용자 목적에 맞게 확장할 수 있다. 이러한 GML의 특성은 통합 GIS데이터 모델로 설계되고 구축된 지형지물을 실제적으로 GIS사업간에 공유하기 위한 데이터 교환 포맷으로 가능성을 가진다. 그리고 다른 포맷의 경우 일단 상용 Tool로 전환을 해야만 데이터에 대한 질의가 가능한 반면 GML은 데이터 추출이 데이터 변환 없이 가능하므로 활용에 있어 효과적인 장점을 가진다.

따라서 본 연구에서는 GML을 기반으로 하여 우리나라의 수치주제도 정보 중에서 지질도와 기타 관련 정보를 구축하고, 운영할 수 있는 실제 사례를 개발하고자 하였다

2. 본론

2.1 GML의 주요 내용

GML 1.0은 XML 1.0을 따르고 있기 때문에 DTD(Document Type Definition)를 기반으로 하고 있으며 OGC의 simple feature를 위한 인코딩 방법을 제시하고 있다. GML 1.0은 다양한 사용자를 위해 세 가지 프로파일로 제시가 되는데 각각은 다음과

같다.

프로파일 1- DTD에 기반을 둔 솔루션을 사용하기를 원하나 응용 DTD 개발이 준비되지 않은 경우나, DTD의 고정된 set에 대해 데이터가 반환되기를 원하는 경우 사용되며, Geometry DTD, Feature DTD가 요구된다.

프로파일 2- DTD에 기반을 둔 솔루션을 사용하기를 원하며 사용자가 응용 DTD 개발이 준비되어 있거나, 참조된 DTD로 인코딩된 데이터를 수용할 수 있게 준비된 경우 사용되며, Geometry DTD와 사용자가 직접 정의한 응용 DTD가 요구된다.

프로파일 3- RDF(Resource Description Format)와 RDF 스키마를 사용하기 위한 경우에 사용된다.

DTD는 텍스트라는 오직 한 가지 데이터 타입만 존재하는데 이는 정수, 실수 등과 같은 데이터 타입에 대한 유효성을 검증할 수 없는 문제점이 있다. 또 DTD에서 네임스페이스를 사용하려면 각각의 네임스페이스에서 모든 요소들을 포함하여야 하는 종속성을 갖는다.

이러한 문제점들은 GML 2.0에서 XML 스키마 언어를 사용하여 해결하였다. 추상적인 데이터 셋의 구조를 설명하는 스키마는 DTD를 보다 발전시킨 것으로 확장성과 재사용성 모듈화에 아주 뛰어나다. 스키마는 XML로 만들어져 있으며 DOM(Document Object Model)을 사용할 수 있으며 DTD의 문제점이었던 데이터 타입과 네임스페이스를 지원하는 특징을 가지고 있다.

GML 2.0 역시 simple feature만을 위한 것으로 기본 기하 태그, 공통 데이터 모델, 그리고 응용 스키마의 생성과 공유를 위한 하나의 메커니즘을 제공함으로써 상호 운용성을 지원한다. 데이터베이스 구현은 지원하는 지리적 피쳐의 스키마 표현들을 정의하는 응용스키마 제공이 필요하며 응용스키마를 만들기 위하여 GML 2.0에서 제공하는 세 가지 기본스키마(Geometry 스키마, Feature 스키마, Xlink 스키마)의 정의들을 사용한다. Geometry 스키마는 추상적인 기하요소와 점, 선, 면 등의 단순 기하객체

모델의 기하객체에 대한 정의를 포함하여, Feature 스키마는 지형지물 식별자, 이름 등의 공통 속성과 지형지물 집합 (Feature Collection) 등 지형지물 유형에 대한 정의를 포함하고 있다. 그리고 Xlink 스키마는 링크의 기능을 제공한다.

이 세 가지의 기본스키마는 사용자들에 의해 바로 사용되어 지는 것이 아니라 사용자들이 각각의 응용 분야에 맞는 응용 스키마를 작성하기 위해 기본 재료로 사용되며 GML 2.0에서는 응용 스키마를 작성하기 위한 구체적인 규칙들을 제시하고 있다.

2002년에 발표된 GML 3.0은 이전 버전에서는 다루지 않았던 여러 기하 요소들과 complex objects를 다루고 있다. GML 3.0에서 새로 추가된 요소들로는 복잡하고 비선형적인 지형지물, 3차원 기하, 2차원 위상구조를 가지는 지형지물, 시간 특성을 가지는 지형지물, 동적인 지형지물, 커버리지, 관측데이터 등이 있다. 그 양이 이전 버전의 7~8에 이를 정도로 방대해 졌으나 사용자는 응용 도메인에 필요한 요소들만 선택적으로 사용하여 응용 스키마를 작성하면 된다. 또한 GML 3.0은 이전 버전과의 높은 호환성을 지원하고 있어 삭제된 몇 가지의 특성을 새로 대체된 요소들로 교환해 주는 정도의 작업만으로 쉽게 업데이트 할 수 있다.

2.2 지질정보의 GML 변환

특정 분야의 GIS 주제정보를 GML로 변환하는 것은 데이터 모델의 설계에 따라 달라지기 때문에 UML로 모델을 설계하여 이에 따라 GML 2를 기반으로 하여 응용스키마를 작성하였다. 이에 대한 구체적인 응용 사례를 보이기 위해 기본 포맷인 shape file의 수치 지질도를 스키마 구조에 따라 인코딩한 뒤, SVG로 렌더링 하는 프로그램을 시험적으로 개발하였으며, 기본 지질 자원 정보로 사용될 수 있는 항공 잔여 자력도 또한 GML로 변환하여 GML을 기타의 특정 응용분야에 적용하기 위한 접근방식을 제시하고자 하였다.

수치 지질도를 GML로 변환하기 위해서는 어떠한 요소들을 어떤 형식으로 표현할 것인지에 대한 기본 모델이 제시되어야 한다. 수치지질도의 데이터 모델은 USGS의 Digital geologic map data model을 참조하여 디자인되었다. 이는 최초로 지질도에 대한 표준을 제시한 것으로, 지질도를 시·공간상에 위치하는 지리개체를 특정한 목적을 위해 기호화하고 그 특징을 기술하는 것으로 정의하였다. 모델의 구체적인 사항들은 USGS의 모델을 우리나라의 수치 지질도와 비교하여 필요한 요소들만 발췌하여 구성되었으며, 지리 정보의 내용은 국립지리원의 수치 지질도 제작 방침을 참고하였다. 본 연구에서 디자인된 데이터 모델이 그림 1에 UML기법으로 표현되었다.

UML은 표준화된 모델링 언어로 높은 호환성과 활용성 때문에 GIS 분야에서 산업계 표준으로 사용되고 있다. UML에서 클래스는 다른 객체들에 대해 공통의 속성, 행동, 관계 그리고 공통의 의미를 갖는 객체의 그룹으로 사각형으로 표시된다. UML 클래스 다이어그램은 모델을 이루고 있는 모든 클래스들을 나타내고 그 사이의 관계를 연결선으로 나타낸 것이다.

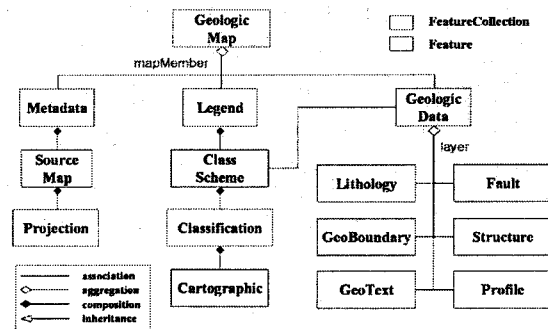


그림 1. 지질도 UML 다이어그램.

GML에서 실세계의 표현은 피쳐들의 집합으로 표현된다. 각 피쳐는 {name, type, value}로 이루어진 프로퍼티들의 집합으로 정의되며, 이들 피쳐들의 집합은 피쳐 컬렉션으로 표현되며, 피쳐 컬렉션과 피쳐들과의 관계는 featureMember라는 연관관계를 통해 나타낸다.

그림 1에서 보는 바와 같이, 최상위 요소인 지질도는 GML의 추상 피쳐 컬렉션 클래스로부터 상속되었으며, 메타데이터(Metadata), 범례(Legend) 그리고 지질 데이터(GeologicData)를 구성 요소로 갖는다. 수치 지질도의 정보 내용 및 기하 특성들은 각 레이어별로 지질데이터 클래스의 구성요소들로 표현되었다.

본 연구에서 메타 데이터는 수치 지질도의 디지털 자료에 포함되는 정보의 소스에 대한 정보를 의미하는 것으로, USGS의 "Digital map data model"에서 정의한 관계형 모델의 메타 데이터 부분을 참고하여 정의하였다. 메타 데이터는 자료의 소스에 대한 일반적인 정보를 포함하는 부분과 소스에 적용된 지도의 투영법을 전체적으로 서술하기 위한 부분으로 이루어져 있다. 지도에서 다루고 있는 지역의 범위, 투영법, 좌표계, 축척 등 사용자가 필요한 데이터를 검색하거나, 데이터가 요구 내용에 맞는 지 확인하기 위해서 메타데이터의 요소들을 이용할 수 있다. 이는 이후에 국가 메타데이터 표준에 맞춰 수정 및 보완이 가능할 것이다.

그림 2는 본 프로그램에서 메타데이터 부분을 GML로 작성하기 위한 개발한 프레임이다. 좌측부에 메타데이터의 전체 구조가 나열되어 있으며 오른쪽 부분에 값들을 입력하면 자동으로 GML문서를 생성한다.

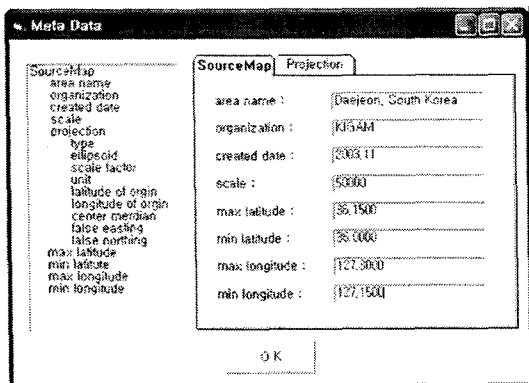


그림 2. 메타데이터 프레임.

범례는 지도에 나타날 요소들을 선택하고 그에 대한 스타일을 정의하는 부분이다. 해당 지리 요소가 속하는 분류 그룹을 지정하

고 각각의 요소에 대한 라벨과 지도에서 표시되는 방법, 범례에서 나타나는 순서 등을 제시하여 사용자가 지도의 내용을 시각적으로 확인할 수 있게 한다.

실제로 수치 지질도가 GML로 변환되는 과정은 Geologic Data이다.

Geologic Data에 대한 보다 상세한 데이터 모델이 그림 3에 나타나 있다.

지질도의 각 레이어들은 지질정보에 대한 모든 필드들을 비공간 속성으로 공간 정보를 기하 속성으로 갖는 피쳐들로 표현되었으며 그림 3에 지질정보에 대한 데이터 모델을 UML 클래스 다이어그램으로 나타내었다. 공간적 분포를 보이고 있는 lithology 피쳐는 폴리곤으로, 특정 위치에 대한 정보인 지질문자는 점으로, 나머지 지질 경계, 단층, 지질 구조, 지질단면선은 선 구조로 표현되었다.

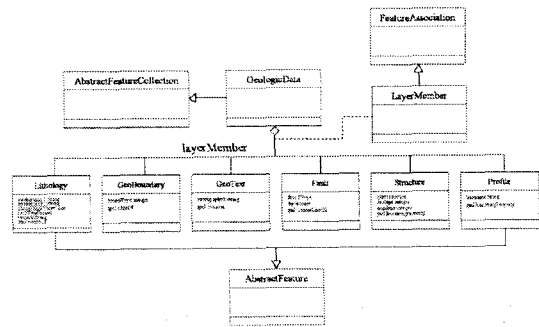


그림 3. 지질정보 UML 다이어그램.

본 프로그램에서는 이번 연구에서 개발된 스키마에 따라 Shapefile 구조의 수치지질도를 자동으로 변환해준다. 먼저 변환할 레이어의 종류를 선택하면 오른쪽에 그 레이어가 가지는 프로퍼티들의 이름이 나열되고, 선택한 파일의 dbf 파일에 있는 모든 필드의 이름이 콤보박스에 표시된다. 사용자는 각각의 프로퍼티로 변환하고자 하는 필드를 선택하면 스키마를 따르는 GML문서가 자동으로 생성, 저장된다 (그림 4, 그림 5).

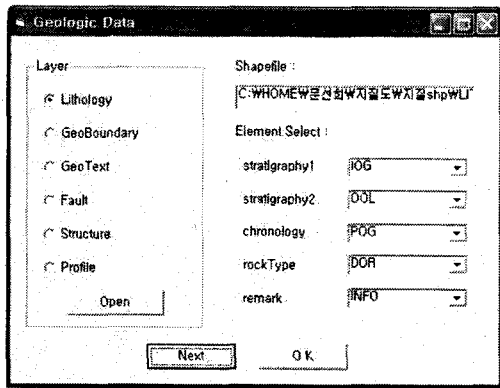


그림 4. 지질 데이터 프레임.

```

<Lithology gml:id="l1">
  <stratigraphy1>Kqv</stratigraphy1>
  <stratigraphy1>Quartz</stratigraphy1>
  <chronology>210</chronology>
  <rockType>Quartz</rockType>
  <gml:extentOf>
    <gml:Polygon gml:id="p1">
      <outerBoundaryIs>
        <LinearRing>
          <gml:coordinates>226857,313174 226883,313067
            226821,313165 226857,313174</gml:coordinates>
        </LinearRing>
      </outerBoundaryIs>
    </gml:Polygon>
  </gml:extentOf>
</Lithology>
  
```

그림 5. GML 문서의 예.

생성된 GML문서는 정보 내용 자체만을 표현하기 때문에 시각화(visualization) 또는 표현(representation)을 위해서는 그래픽 포맷으로 변환되어야 한다. 이번 연구에서는 시각화를 위해서 XML을 기반으로 하고 있는 SVG(Scalable Vector Graphic)를 이용하였다. GML을 SVG로 표현하기 위해서는 어떤 정보의 내용을 어떻게 표현할 것인지 스타일을 정의해 주어야 하는데 XSL(XML Stylesheet Language)를 이용하였다. GML이 내용과 표현을 분리하는 이러한 특성으로 인해 하나의 지질정보 GML에 대해 사용자의 요구나 작업하고자 하는 내용에 따라 다양하게 표현할 수 있다.

그림 6은 같은 GML 문서에 대해 서로 다르게 스타일링 한 예시 결과를 보여준다. 현재 SVG는 영어만을 지원하고 있기 때문에 한글이 포함된 수치 지질도의 파일 내용을 그대로 GML로 변환하고 이를 랜더링하는 것은 지원되지 않는다. 따라서 본 연구에서는 한글로 작성된 층서는 영어 약자로

대체하였으며, 기타정보를 담고 있는 remark 요소는 영어로 바꿔서 입력하였다.

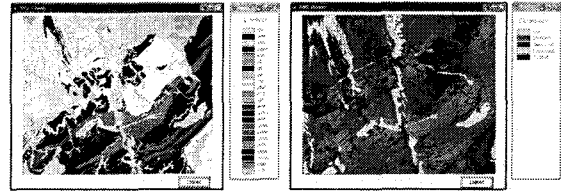


그림 6. (a) 층서에 따른 시각화
(b) 지질시대에 따른 시각화.

또한 본 연구에서는 항공자력탐사 자료를 통해 얻은 잔여 자력값을 지질도와 함께 시각화하였다. 관측점의 위치를 중심으로 그리드의 간격을 크기로 하는 셀을 지정하고 잔여 자력값에 따라 점이적으로 색을 변화하여 나타내었다.

지구에 대해 상대적인 위치를 갖는 모든 정보들은 그 좌표를 지질도의 좌표체계인 KTM(Korean Transverse Mercator) 좌표계로 지정해주면 어떤 자료든지 같이 웹 브라우저 상에 나타내어 통합 분석할 수 있다.

그림 7은 제주도 지역의 지질도와 항공 잔여 자력도를 GML로 변환하여 SVG 지도로 랜더링한 결과이다.

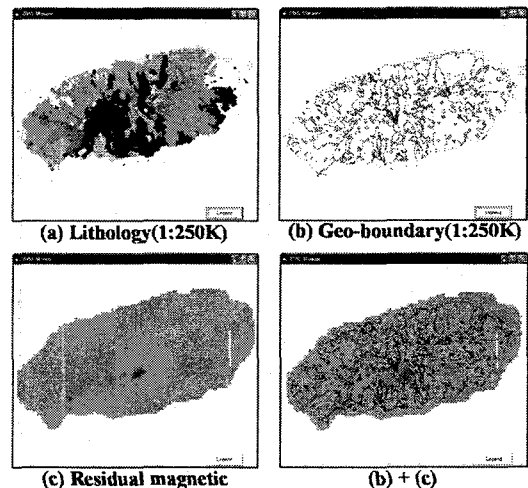


그림 7. (a) 기본지질정도 (b) 지질경계
(c) 잔여자력도

3. 결론

본 연구에서 개발된 프로그램은 지질 분야에서 기존에 작성되어 있는 자료들을 GML로 변환하기 위한 기능을 구현하고 있다. 생성된 GML 데이터는 웹 상에서 별도의 프로그램 없이 확인이 가능하며, DOM을 통해 피처를 개별적으로 접근할 수 있어 GIS tool에서와 같은 공간 분석 및 검색이 가능할 것이다. 피처를 클릭했을 때 일어나는 피처기반 질의는 기존의 래스터 이미지 형식에서는 가능하지 않으며, 구현한다 하더라도 픽셀 위치를 파악하여 지리좌표로 변환하고 다시 서버에서 해당하는 피처를 찾는 복잡한 과정을 거치게 된다. 현재까지는 지리정보시스템 데이터가 생산자 중심의 구축에 집중되어 왔던 반면, GML은 사용자가 데이터를 쉽게 접근할 수 있고, 사용자가 원하는 데이터를 획득할 수 있어 사용자 관점에서의 지리정보 시스템 데이터의 활용성을 높이고자 하는 적극적인 노력의 하나라고 하겠다. 이와 같은 GML의 특징들은 GIS 툴에 익숙하지 않은 일반 사용자들의 지질 데이터 사용에 큰 이점을 제공할 것으로 생각된다.

GIS의 각 응용 분야에서 GML 응용 스키마를 개발하고 공유함으로써 데이터 제작의 표준을 이룰 수 있으며, 이에 따라 본 연구에서 이루어진 지질 데이터와 동일한 방법으로 GML변환이 가능할 것이다. 같은 스키마에 의해 생성된 GML 문서는 같은 데이터 구조를 가지므로 하나의 stylesheet 문서만으로 많은 문서들을 동일한 규칙으로 SVG 렌더링이 가능하다. 또한 이와는 반대로, 하나의 GML 문서에 대해서 원하는 데이터만 추출하여 시각화하거나 사용자의 요구 또는 분석의 목적에 따라 서로 다른 방법으

로 시각화하는 것도 가능하다.

Polygon 형태를 변환한 GML 문서의 경우, 그 좌표의 수로 인해 shape file에 비해 큰 용량으로 작성되었지만, 반복되는 구조의 텍스트 파일이기 때문에 일반 파일 압축을 통해서도 20%이하로 줄어들었다. 현재 GML의 전송량을 줄이기 위해 제안되고 있는 gzip, XML 등의 압축 방법을 이용하면 동일한 지리 정보 데이터를 고유의 바이너리 포맷으로 저장할 때에 비해 데이터양이 적은 것으로 보고 된 예가 있다. GML 압축 기술로 인한 데이터 크기의 감소는 웹 브라우저 뿐만 아니라 PDA나 휴대폰 등 전송량에 제한이 많은 이동통신 기기에서 특히 효율적일 것으로 생각된다.

참고문헌

- [1] C. Simon, C. Adrian, Ron, and M. Richard, "Geography Markup Language, ver2.0", Open GIS Consortium, 2001.
- [2] Bruce R. Johnson, Boyan Brodaric, Gary L. Raines, Jordan T. Hastings, and Ron Wahl, " Digital geologic map data model, U.S. Geological Survey", 1999.
- [3] 수치 지질도 제작 방침, 국립지리연구원, 1999.
- [4] Simon W.Houlding, "XML-an opportunity for meaningful data standards in the geoscience", Computers & Geosciences vol.27, 2002, pp.839-849.
- [5] J.Jellema, R.J. van Leeuwen, "Using XML in advanced geological information systems", EAGE, 2002.
- [6] 국가 GIS 통합 데이터 모델 확립 연구, 건설교통부, 2002.