

# 하천정보의 효율적 활용을 위한 메타데이터의 구축에 관한 연구

김계현<sup>1)</sup>, 김태화<sup>2)</sup>

## 1. 연구의 배경 및 목적

세계는 고도기술 사회를 지향하며 정보사회로 빠르게 이전하고 있으며 정보를 국가 재원으로 인식하고 이에 효율적으로 대처하기 위한 노력을 기울이고 있다. 정보의 가치와 유용성은 바로 부가가치의 척도이며 의미 있는 자료는 재원을 확보할 수 있는 잠재력을 가진 것으로 판단되고 있다. 따라서 복잡한 자료의 효과적 처리와 표현은 매우 중요한 일로 인식되고 있으며 이러한 자료를 사용자에게 적합한 자료로 제작하고 공급하는 것은 매우 중요한 일이다. 하천정보의 활용에 있어 방대한 수치공간자료의 관리를 위해 명확하고 체계적인 정보공유의 수단으로 메타데이터의 제공은 필수적인 사항이다. 현재 여러 기관에서 구축되고 있는 다양한 하천정보를 공유한다면 정보의 중복을 피할 수 있으며 그로 인한 경제적인 효과를 기대할 수 있다.

본 연구에서는 국가 표준의 하천정보 관리 및 원활한 공급과 활용을 위한 것으로 하천정보의 생산과정, 품질, 기준, 공급 등을 명시한 메타데이터 설계를 목적으로 한다. 메타데이터의 설계는 사용자 및 관리자가 하천정보를 활용하는데 있어 필요한 정보를 추상화, 목록화 하는 것을 말한다. 따라서 설계를 위해 국외의 메타데이터 현황을 조사하여 메타데이터의 범위와 구성을 파악하였으며, 하천정보 메타데이터의 설계를 위하여 하천 관련 자료를 조사·분석하였으며 메타데이터의 설계시 주요 고려사항을 제시하였다.

## 2. 하천정보 표준 메타데이터의 정의

메타데이터란 수록된 자료의 내용, 품질, 조건 및 특징 등을 저장한 자료로서 자료의 이력서라 할 수 있다. 메타데이터는 일반적으로 자료의 양이 방대하거나, 자료의 직접 접근이 용이하지 않을 경우 자료를 참조하기 위한 자료로 이용되며, 메타데이터의 대한 질의를 통하여 지리정보의 취득이 가능하도록 인터넷 환경에서 지리정보의 공급이 가능토록 한다. 메타데이터의 주요 기능은 작성한 실무자가 바뀌더라도 변함없는 데이터의 기본체계를 유지하여 일정 시간이 경과하여도 일관성 있는 데이터를 이용자에게 제공할 수 있다는 점이다. 또한, 표준화를 통한 정보공유의 극대화를 도모하며 데이터의 원활한 교환을 지원하기 위한 틀을 제공함에 있다.

하천정보의 범주는 하천제방도, 하천부속물도, 하천종횡단도, 하천지적도, 하천자원도, 하천주변지형지물도, 수문관측망도 등이며 이러한 자료의 이력 및 내용, 품질, 공급 등을 제공하는 것을 하천정보 표준 메타데이터라 칭한다.

1) 인하대학교 지리정보공학과 부교수·공학박사·032-860-7602(E-mail:kyehyun@inha.ac.kr)

2) 인하대학교 지리정보공학과 석사과정·032-860-4413(E-mail:g1991676@inhavision.inha.ac.kr)

### 3. 메타데이터 표준 모델

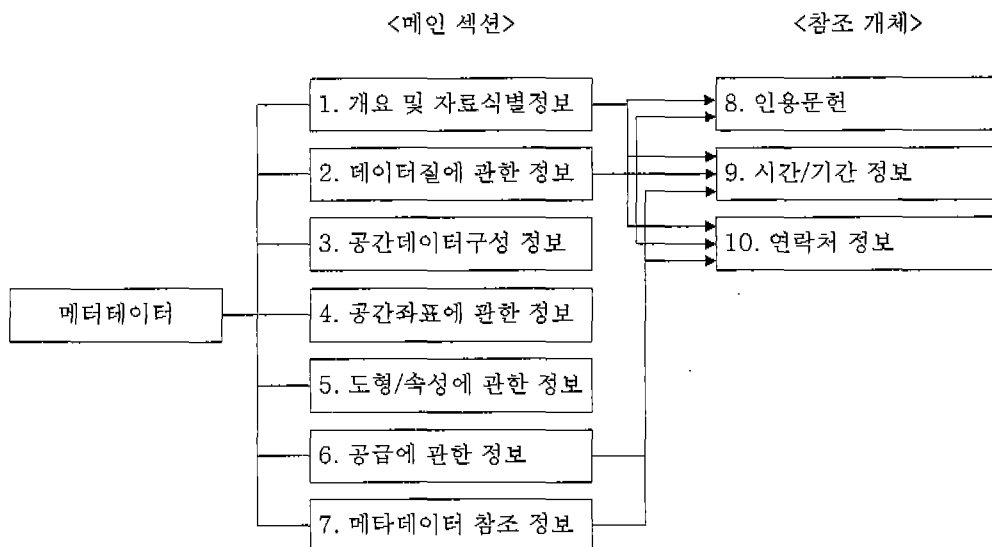
해외 표준화 기구의 사례와 함께 미국의 FGDC, 국제표준화기구와 같은 선진 사례를 분석하여 효율성 높은 메타데이터의 구축에 반영하였다. 특히 선진외국의 사례 분석에서 국가별 메타데이터의 특성과 차별화 되는 사항을 심층 분석하여 국내의 메타데이터의 설계에 보다 효율성을 부여하도록 하였다.

#### 3.1 미국 FGDC 메타데이터

FGDC(Federal Geographic Data Committee) 메타데이터의 주요 사용목적은 세가지로 요약된다. 첫 번째는 GIS 데이터를 생성하는 기관의 내부적인 데이터 관리에 활용되는 것과 두 번째는 GIS 데이터 관련 공급기관에 데이터 생성 기관에서 보유하고 있는 데이터에 대한 정보를 제공하는 것이다. 세 번째는 GIS 데이터의 처리 및 가공이 필요한 데이터에 대한 정보를 제공하는데 그 목적이 있다. FGDC의 메타표준은 메타데이터에 관한 용어상의 정의와 활용에 있어 표준을 제시하여 사용의 통일성을 기하기 위하여 제정되었다. 메타데이터의 주요 내용은 자료가 담고있는 내용, 자료의 품질, 상태 등에 관한 사항이다.

FGDC 메타데이터 표준은 단지 지형공간 메타데이터 (Geospatial Metadata)의 논리적 구조와 내용만을 표준화한 내용 표준을 의미하며, 관계형 데이터 베이스 내에서의 운영과 같은 구체적인 모형을 제공하지는 않는다. 반면, 각기 다른 플랫폼에 의해 생성된 자료의 탐색 및 상호 호환성을 구축하기 위해서는 교환 포맷에 대한 규칙은 FGDC의 공간정보 유통기구인 Clearinghouse의 XML 교환 표준을 사용하고 있다.

FGDC의 메타데이터는 11개의 장으로 구성되어 있다. 7개의 주요 장과 3개의 보조 장으로 구성되며<그림 1>, 7개의 주요 장은 필수항목(Mandatory)과 조건부 필수항목(Mandatory if applicable)으로 나눈다.



<그림 1> FGDC 메타데이터 구성도

### 3.2 국제표준화기구의 메타데이터

ISO는 표준 관련 국제기구이다. 관련 기술위원회는 96년에 메타데이터 초안을 발표한 이래 97년 1월에 버전 2.0을 내놓았고, 97년 3월에는 버전 3.0 초안과 98년 5월에 버전 4.4를 발표하였다. 현재는 버전 4.5 까지 발표하였으며, 버전 4.5에서는 적합성 레벨을 두어 지리정보의 기술에 있어 차등을 두었다. 데이터의 식별, 평가, 사용관리에 대한 관한 정보는 다음<표 1>와 같다.

<표 1> ISO 메타데이터 표준의 주요 내용

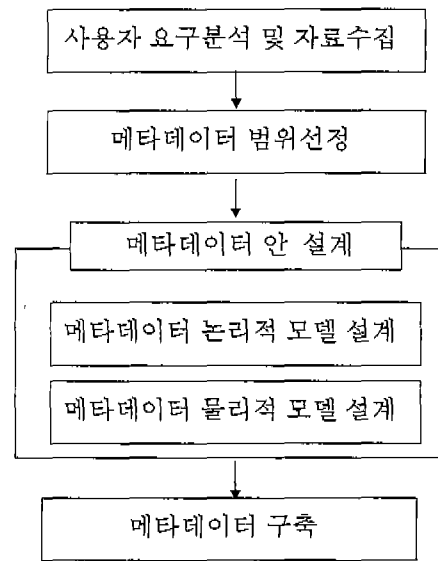
항 목	설 명
1.식별정보 (Identification)	사용언어, 요약, 목적, 부가설명, 데이터 셋 환경, 식별인용, 데이터 셋 인증, 착수업무 실별정보, 이미지정보, 상태정보, 데이터 셋 범위, 지역범위, 시간적 범위, 수직범위, 분류, 주요연락처, 도식용 그래픽, 데이터 셋 관계 데이터 셋 제약
2.품질정보 (Data quality)	데이터 품질 적용대상, 데이터 품질 적용대상 설명, 데이터 품질 지리적 범위, 데이터 품질 시간적 범위, 데이터 품질 평가, 정량적 평가, 정량적 평가참조
3.연혁정보 (Lineage)	연혁적용대상, 연혁적용대상 설명, 연혁 지리적 범위, 연혁 시간적 범위, 연혁보고, 진행단계, 기타용도
4.공간데이터표현정보 (Spatial data representation)	공간정보표현방식, 벡터공간표현정보, 레스터공간표현정보, 영상공간표현정보, 감지밴드정보
5.기준체계정보 (Reference systems)	시간기준체계명, 시간기준체계 참고사항, 시간기준체계설명, 달력정보, 시간좌표계정보, 순서적인 시간기준체계정보, 순서적인 연대정보, 지리적 식별자에 의한 공간기준체계, 좌표에 의한 공간참조, 좌표기준체계 참고사항, 좌표기준체계 다른 이름, 좌표기준체계 기준일, 데이터, 지구 타원체, 좌표체계, 좌표변환, 좌표연혁, 좌표변환 정확도
6.지형지물 목록정보 (Feature catalogue)	외부지형지물 목록, 내부지형지물 목록, 지형지물 목록생산자, 지형지물 유형
7.자료배포정보 (Distribution)	배급자정보(배포식별자, 배포책임, 고객주문과정, 표준주문절차), 배포포맷정보(배포포맷명, 배포되는 정보내용, 배포 온라인 지원, 기록용량 등)
8.메타데이터 참조정보 (Metadata reference)	메타데이터 파일 식별자, 메타데이터 상위 식별자, 데이터 셋 응용스키마 명, 메타데이터 표준정보, 메타데이터 부가정보, 메타데이터 제약사항, 메타데이터 보안사항, 메타데이터 확장정보
9.참고자료(Citation)	제목, 부제목, 참고자료 책임담당자, 기준날짜, 기준날짜 유형, 판, 데이터 셋 식별자, 데이터 셋 유형, 표현방식, 시리즈 간행물, 간행물 발행번호 식별정보, 다른 참조 자료 정보
10.책임담당자정보 (Responsible Party)	책임자 개인명, 책임기관, 책임담당자 직위명, 책임담당자 역할, 책임담당자 주소
11.주소정보 (Address)	세부주소, 도시, 행정구역, 우편번호, 국가, 음성전화, TDD/TTY 전화, 팩스전화, 전자우편 주소, 네트워크 주소, 서비스 기간, 연락처침

#### 4. 하천정보 메타데이터 설계

하천정보 분야의 메타데이터 설계의 주요 관점은 국가기관이 보유하고 있는 하천 관련 정보의 활용을 극대화하기 위한 표준화를 지향할 수 있어야 한다. 즉 국가기관의 하천관리 업무를 조사·분석하여 분류한 내용과 표준화 작업에 초점을 맞추고 제작된 메타데이터를 분석하여 두 결과를 접목시켜 설계에 반영하여야 한다.

##### 4.1. 메타데이터 설계 절차

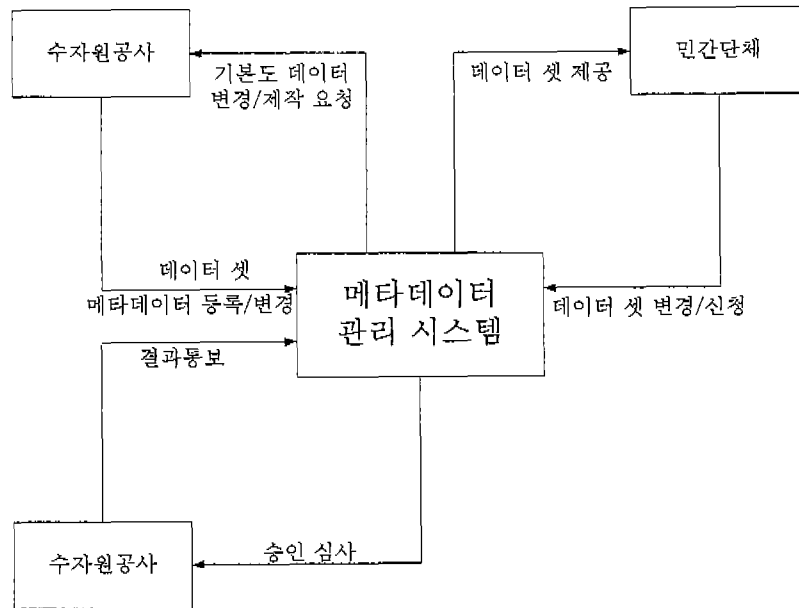
설계를 위한 작업 단계 및 분석과정은 사용자 요구분석 및 자료수집, 메타데이터 범위선정, 메타데이터(안) 설계, 메타데이터 구축으로 분류할 수 있다. 먼저, 사용자 요구분석 및 자료수집 단계에서는 메타데이터의 기술동향 및 구축사례를 파악하며, 유역관련 업무분석을 통한 메타데이터 관련 업무를 도출한다. 범위선정 단계에서는 국내외 메타데이터 표준화를 고려하여 하천정보 관리에 적합한 메타데이터 표준 모델을 선정한다. 메타데이터(안) 설계 단계에서는 하천정보의 이력을 제공하는 메타데이터 항목을 선정하는 단계이다. 이러한 항목선정의 판단기준은 하천정보 관리업무의 조사·분석을 통하여 이루어지며, 항목의 중요도에 따라 필수 항목, 조건부 항목, 선택항목으로 분류되어지게 된다. 메타데이터 구성항목이 선정되면 이를 논리적 모델 및 물리적 모델로 설계한다. 메타데이터 구축 단계에서는 메타데이터 설계안에 따라 대상지역을 선정하여 메타데이터 설계(안)의 검증에 위한 메타데이터 시범구축을 실시한다<그림 2>.



<그림 2> 메타데이터 설계 절차

##### 4.2 사용자 요구분석 및 자료수집

하천정보는 유역특성 및 유역특성인자와 유역형상 및 하천 종단도와 같은 다양한 정보가 존재하며, 이러한 정보의 추출과정을 분석하여 메타데이터 설계과정에 반영한다. 또한 메타데이터 적용시 업무변화를 고려한다. 메타데이터의 관리업무는 데이터 셋의 제작, 승인, 공급 등의 주요 세 가지 업무로 분류되며 각 작업은 다음과 같은 관계를 가진다(그림 3).



<그림 3> 메타데이터 관리 업무

#### 4.3 메타데이터 범위선정

‘국가지리정보체계(NGIS) 정보유통을 위한 정보기록 방식(메타데이터) 표준화를 위한 연구’에서 우리 나라는 ISO/TC211의 메타데이터를 채택하여 이용하는 것이 여러 측면에서 타당하다고 판단하였다. 우리 나라의 경우 각 기관에서 메타데이터의 연구와 실제적인 구축사례가 있으며, 이러한 구축사례는 ISO/TC211 표준에 근거를 두고 있다. 따라서 기존의 메타데이터의 범위와 현재 유역정보관리의 행정업무적인 상황과 수치지도화에 따른 제약사항 등을 반영한다. 이를 위해 유역정보 관련 업무의 구성요소와 연계 작업을 분석하였다. 분석과정은 업무환경다이어그램, 업무흐름도, 자료 흐름도의 기록방식을 사용하여 문서화되도록 하였다. 메타데이터의 설계시 고려되어야 할 사항을 확인한 후 메타데이터 구축을 위한 메타데이터 기준을 설정하며, 최종적으로 메타데이터 관리시스템의 설계범위를 결정하였다.

#### 5. 메타데이터의 설계

사용자 요구를 최대한 반영하기 위하여 필요시 기관내부의 사용을 위한 내부메타데이터와 외부 사용자를 위한 외부메타데이터로 나누어 설계를 실시하되, 가능한 국제 메타데이터 표준안에 부합하도록 통합안을 지향하였다. 메타데이터의 구성요소가 결정되면 다음의 설계순서에 따라 메타데이터를 설계하였다.

우선적으로 하천정보 메타데이터 설계안은 유역정보와 배포를 위한 정보를 정의하였다. 아울러 유역정보의 효율적 제공을 통한 광범위한 사용자의 목적에 부합토록 하기 위하여 전반적인 하천자료의 메타데이터를 설계하고 구축하였다. 이와 함께 사용자의 요구사항, 국내외 표준 동향, 하천정보의 특성 등을 고려하여 메타데이터 설계안을 제시하였다. 제시된 설계안에 대하여는 관련 사용자와 전문가의 의견을 수렴하여 효율성을 제고하였다. 메타데이터의 구성은 하천정보에 대한 목

록, 식별, 품질, 연혁, 공간 데이터 표현 및 기준체계, 지형지물목록, 자료배포, 메타데이터참조, 참고자료, 책임담당자, 주소 등의 정보를 제공하도록 고려되었다 <표 2>.

<표 2> 하천정보 메타데이터 표준의 주요 내용

항 목	설 명
0.목록정보	데이터의 내용, 질, 상태, 입수 방법, 특성등 전반적인 사항을 기술
1.식별정보	데이터에 대한 기본적이고 개괄적인 정보를 기술. 주요항목들은 제목, 목적, 지리적 영역, 현재성, 데이터의 상태, 대표어, 보안적인 사항, 데이터의 입수, 이용을 위한 사항을 포함
2.품질정보	데이터의 품질에 관련된 정보를 기술. 주요항목들은 데이터의 품질 평가, 위치 및 속성의 정확성, 완결성, 일치성, 데이터에 대한 원 자료 정보, 데이터의 생산방법 등을 포함
3.연혁정보	연혁적용대상, 연혁적용대상 설명, 연혁 지리적 범위, 연혁 시간적 범위, 연혁보고, 진행단계, 기타 용도 등을 포함
4.공간데이터표현정보	데이터의 공간정보를 표현하는데 사용된 방법을 기술. 이를 위하여 직접 혹은 간접적인 방법으로 표현하고 있으며, 공간 객체(Object)로 표현한다. 주요내용은 간접 공간참조 자료(주소체계), 직접 공간참조 자료(벡터 혹은 래스터), 객체에 의한 표현자료(점, 선, 면의 개수), 위상관계를 포함
5.기준체계정보	공간정보 참조를 위한 방법과 인코딩 수단, 좌표계에 대해 기술. 주요항목은 지도 투영법 및 직교 좌표체계의 이름, 수평 및 수직 해상도, 좌표계 해상도를 포함
6.지형지물 목록정보	외부지형지물 목록, 내부지형지물 목록, 지형지물 목록생산자, 지형지물 유형
7.자료배포정보	데이터의 획득에 관한 방법, 공급자에 관한 정보, 이용 매체 등과 같은 정보를 기술. 주요 항목은 배포자의 연락처, 이용 가능한 포맷, 온라인이나 카트리지 테이프, CD-ROM과 같은 물리적 매체를 통해 자료를 얻는 방법에 대한 정보, 자료취득에 드는 비용을 포함
8.메타데이터 참조정보	메타데이터 파일 식별자, 메타데이터 상위 식별자, 데이터 셋 응용스키마 명, 메타데이터 표준정보, 메타데이터 부가정보, 메타데이터 제약사항, 메타데이터 보안사항, 메타데이터 확장정보 등의 포함
9.참고자료	제목, 부제목, 참고자료 책임담당자, 기준날짜, 기준날짜 유형, 판, 데이터 셋 식별자, 데이터 셋 유형, 표현방식, 시리즈 간행물, 간행물 발행번호 식별정보, 다른 참조 자료 정보 등의 포함
10.책임담당자정보	책임자 개인명, 책임기관, 책임담당자 직위명, 책임담당자 역할, 책임담당자 주소 등의 포함
11.주소정보	세부주소, 도시, 행정구역, 우편번호, 국가, 음성전화, TDD/TTY 전화, 팩스전화, 전자우편 주소, 네트워크 주소, 서비스 기간, 연락처침 등의 포함

## 6. 메타데이터의 시범 구축

선정된 시범 지역에 대한 메타데이터 관련 자료를 수집하여 시범적으로 구축하였다. 가능한 충실한 메타데이터의 시범구축을 위하여 전반적으로 메타자료가 유용한 지역을 시범지역으로 선정하며, 최대한 장기적인 데이터의 확보를 도모하였다. 메타자료가 존재하지 않는 항목에 대하여는 임의의 자료를 구축하여 대체하였다.

유역정보의 구성요소가 되는 하천정보, 수문정보, DEM등의 자료를 시범적으로 제작하였다. 사용자 이용 목적에 적합한 메타데이터를 제공하기 위해 적합성 수준 1과 수준 2로 구분하였으며, 적합성 수준1은 일반적인 자료명세의 제공으로 국한하였다. 적합성 수준 2는 자료의 생산에서 배포까지의 주요 메타데이터 항목을 상세히 기술하였으며, 표 3은 하천도에 대해서 적합성 수준 1로 실제 메타데이터를 설계하였다.

<표 3> 하천도 메타데이터

항 목 명	내 용
목록정보	
제목	북한강 상류 하천도
판	한국수자원공사 하천도 메타데이터 버전1.0
시리즈 간행물명	한강유역조사
간행물 발행번호 식별자	
착수업무 식별정보	
착수업무 유형	연구과제
착수업무 식별명	수자원 관련 공간정보 분석 및 적용에 관한 연구
기준 일자	2001/3/28
책임 담당자 정보	
책임자 개인명	홍길동
책임기관명	한국수자원공사
책임담당자 역할	관리자
세부주소	대덕구 연축동 산6의2
도시	대전광역시
행정구역	대전광역시
우편번호	306-711
국가	대한민국
네트워크주소	<a href="http://www.kowaco.or.kr">www.kowaco.or.kr</a>
전자우편주소	<a href="mailto:kowaco@kowaco.or.kr">kowaco@kowaco.or.kr</a>
음성전화	(042)629-2720

항 목 명	내 용
데이터 셋 지역범위	
지역범위	
서쪽경계좌표	227879.50
동쪽경계좌표	267736.50
북쪽경계좌표	280925.60
남쪽경계좌표	231810.00
지리적범위명	용담댐 상류
지리적 참조사항	수자원단위지도
날짜	2001/3/28
해상도 수준	1: 5000
데이터셋 언어	한국어, 영어
데이터셋 문자코드 집합	
요약설명	수치지형도로부터 하천 레이어를 추출하였다. 제공되는 매체는 CD를 통해 제공되며, 복제를 통한 타인에게 양도는 법적으로 금하여 모든 지적 재산권은 수자원공사에게 귀속된다.
목적	“한강유역조사 사업”은 수자원 장기종합계획 등 유역수자원계획수립에 관련된 기본정보를 제공하고 바람직한 유역의 비전 설정 및 종합관리방안을 제시하는데 그 목적이 있음
진행사항	진행단계
주제	수문학
핵심어	북한강, 하천, 하천도, stream
접근 제한	없음
이용 제한	본 데이터는 연구용으로만 사용할 수 있으며 그 외 상업적 목적으로 사용할 수 없다
연혁	
서술식 품질 평가보고서	
공간적 표현방식	벡터
기준좌표계 유형	좌표
배포 식별자	
배포 포맷명	SHP
배포 매체	CDROM, On-line
배포 온라인 리소스 URL	<a href="http://www.kowaco.or.kr">www.kowaco.or.kr</a>
적합성 수준	1
메타데이터 언어	한국어,영어
메타 데이터 문자코드 셋	
메타 데이터 최종 수정일	2001/3/28



## 7. 메타데이터 구축시 고려사항

메타데이터 설계시 범위 선정의 기준이 되는 것은 유역정보를 제공하는 사용자 계층의 선정과 메타데이터를 관리방법, 효율적인 유통관리 기능이 충분히 고려되어 있는가를 결정하는 것이다. 따라서 사용자의 수요와 적용계층, 하천정보 관리방법의 조사를 위한 설문조사 및 사용자 요구분석을 실시하여야 한다. 또한 메타데이터의 유통관리를 위해 클리어링하우스 개념의 수자원 정보 유통기구 체계에 대한 사항을 고려하여야 한다.

ISO/TC211 메타데이터 표준 모델은 항목이 매우 많고 복잡해서 정확히 이해하기 어려우며 각 항목의 요소의 내용을 기술하는 방법이 표준화 되어있지 않아 사용자에게 따라 그 내용이 변할 수 있다. 이를 위해서는 메타데이터 설계와 함께 메타데이터 시범구축이 병행되어야 한다. 또한 효율적인 메타데이터의 시범구축을 위하여 전반적으로 메타데이터의 자료확보가 유용한 지역을 시범 지역으로 선정하여 장기적인 데이터의 확보를 도모하여야 한다. 메타데이터의 입력자료의 검증을 위한 메타데이터 편집기를 개발할 경우 국가표준에 부합하도록 설계되어야 하며, 메타데이터의 교환을 위한 양식이 정의되어야 한다.

## 8. 참고문헌

- 1) 국립지리원, 지리정보 관리 및 공급시스템에 관한 연구(1), 1998.
- 2) 김계현, 임삼성, 김희두, 외부메타데이터 및 공급시스템에 관한 연구, 인하대학교, 1998.
- 3) 김창호, 진희채, 이소영, 시종익, 강승림, 이성과, 박정철, 국가지리정보체계(NGIS) 정보유통을 위한 정보기록방식(메타데이터)표준화를 위한 연구, 한국전산원, 1977.
- 4) 정문섭, 신각척, 최홍준, 공간정보의 유통 및 활용촉진을 위한 법제도입방안, 국토개발연구원, 1997.
- 6) Lagoze, Carl. "The warwick framework: A container architecture for diverse sets of metadata", D-Lib Magazine, 1996.
- 7) Lange, Holley R. and B. Jean Winkler. "Taming the Internet: Metadata, A work in progress". Advances in Librarianship, vol. 21, 1997.