Korean Journal of Remote Sensing, Vol.26, No.5, 2010, pp.549~559

공간정보 표준 메타데이터 추출 및 변환 프로그램 개발

한선묵 · 이기원†
한성대학교 정보시스템공학과

Program Development for Automatic Extraction and Transformation of Standard Metadata of Geo-spatial Data

Sun Mook Han and Kiwon Lee†
Dept. of Information Systems Engineering, Hansung University

Abstract : In geo-spatial information system building and operation, metadata is one of the crucial factors. Therefore, international and domestic organizations or associations for standardization have developed and distributed geo-based standard metadata to meet public demands. However, because metadata is composed of complicated elements and needs XML storage and management, individual organization which implement and operate practical application system is inclined to define and use its own metadata specifications. In this study, metadata extraction program, that metadata elements are directly extracted from geo-based file formats was developed to easily utilize standard metadata such as ISO/TC 19115, TTAS.KO-10.0139 and TTAS.IS-19115, and those elements are processed into XML. Furthermore, geo-based images sets are applied to another metadata of ISO/TC 19115-2. As well, metadata transformation is needed due to inconsistent or non-corresponding definition among standard metadata; in this program, transformation modules are also implemented to interoperable uses between standard metadata specifications. Widely used data formats are dealt with in this program, but extension for other formats and other metadata specifications is possible, and it is expected that availability of standard metadata is increased, through this kind of development.

Key Words : ISO/TC 19115, TTA Standard metadata, Extraction, Transformation, Geo-based images.

요약 : 공간정보 시스템의 구축과 운영에서 메타데이터는 핵심 요소 중의 하나이다. 따라서 국내외 표준화 기구나 단체에서 현실적인 수요를 반영한 다양한 메타데이터를 개발, 공표하고 있다. 그러나 국제 표준이나 국가 표준 등이 많은 항목으로 구성되어 있고 메타데이터 서버를 구축하는 경우에는 XML 형식으로 저장, 관리할 필요가 있기 때문에 실무 응용 시스템에서는 표준 형식을 따르기 보다는 운영 기관별로 별도의 메타데이터 사양을 설정하여 사용하는 경우가 많다. 따라서 본 연구에서는 국제 표준인 ISO/TC 19115와 TTAS.KO-10.0139와 TTAS.IS-19115와 같은 단체 메타데이터 표준으로 손쉽게 입력할 수 있도록 직접 공간자료 포맷으로부터 메타데이터 항목과 요소를 추출하여 직접 XML로 저장할 수 있는 프로그램을 개발하였다. 본 프로그램에서 공간영상정보에 대한 자료 포맷의 경우에는 ISO/TC 19115-2 메타데이터 표준 사양으로 추출, 저장할 수 있도록 하였다. 한편 메타데이터 표준간의 항목이나 요소들이 일부 상이하기 때문에

접수일(2010년 9월 30일), 수정일(1차 : 2010년 10월 22일), 게재확정일(2010년 10월 23일).
† 교신저자: 이기원(kilee@hansung.ac.kr)

메타데이터 추출 기능과 함께 메타데이터 간의 변환 기능을 개발하여 메타데이터의 상호 운영이 가능하도록 하였다. 본 연구에서는 활용도가 높은 자료 포맷을 대상으로 하였으나 기타 자료 포맷이나 기관별 메타데이터 사양을 처리할 수 있는 확장 기능 구현을 통하여 메타데이터 표준의 활용도를 증가시킬 수 있을 것으로 기대한다.

1. 서 론

정보자원을 다루는 대부분의 정보시스템에서 메타데이터는 중요한 정보이다. 메타데이터가 '데이터에 관한 데이터' 로서 정보자원의 속성을 기술하는 데이터라는 것은 보편적으로 인식되는 정의이며, 이는 실 데이터와 직접적으로 혹은 간접적으로 연관된 정보를 제공하는 데이터를 의미한다. 메타데이터를 사용하면, 사용자가 원하는 데이터가 맞는가를 확인할 수 있고, 쉽고 빠르게 원하는 데이터를 찾아낼 수 있다.

따라서 메타데이터는 '정보자원을 설명하는 요소의 집합' 으로 정의할 수 있다. 데이터 관리 측면에서도 메타데이터는 자료의 이력 정보를 포함하고 있으므로 자료의 갱신이나 재생산에 이용될 수 있다. 즉, 자료의 계속 유지, 갱신, 삭제되어야 하는 지를 결정하는 정보로 이용될뿐더러 자료에 대한 신뢰도를 판단하는 근거로 이용된다는 것을 의미한다. 한편 프로젝트 관리 측면에서도 메타데이터는 자료의 유형과 내용을 문서화하고 계획하는데 이용되고 자료 구축의 모니터링 과정을 알 수 있도록 한다.

공간정보 메타데이터의 경우도 이러한 일반적인 개념과 활용 목적을 벗어나지는 않는다. 이와 함께 공간정보 메타데이터는 복잡하고 다양한 유형으로 이루어진 공간자료의 시스템 내부적인 구조화(박용재와 이기원, 2010), 카탈로그 서비스(한선묵과 이기원, 2010), 데이터 마트나 데이터 웨어하우스의 운영에 적용할 수 있도록 시스템 구성 자료에 대한 정보를 제공하고, 어떤 변환 과정을 통하여 외부로의 서비스로부터 입력된 자료의 처리와 해석을 위한 정보를 제공한다. 즉, 데이터를 소유하고 있는 측면에서는 관리의 용이성을, 데이터를 사용하고 있는 측면에서는 검색의 용이성을 보장받을 수 있기 때문에 메타데이터의 필요성이 더욱 높아지고 있다(Baca, 2008).

한편 정보통신기술 측면에서도 메타데이터의 중요성은 점점 증가하고 있다(Lawrence *et al.*, 2008). 서비스 기반 어플리케이션을 개발하거나 운용하는 경우 어떤 비즈니스 프로세스에 대해서 어떠한 XML/SOAP (Simple Object Access Protocol) 서비스가 존재하는지, 어떤 컴포넌트를 호출하는지, 어떤 데이터베이스에 있는 어떤 데이터 요소를 사용하는 지 등과 같은 정보는 중요한 요소가 된다. 예를 들어, 다중의 사용자가 데이터 웨어하우스를 공유하기 위해서는 추적해야 하는 데이터 요소에 대한 합의를 해야 한다. 데이터 웨어하우스의 최종 사용자들은 데이터 요소의 출처와 관심 있는 데이터 요소에 대하여 어떠한 질의에 의하여 데이터 자원이 어떻게 변화하였는지에 대한 정보를 필요로 하는 경우도 있다. 따라서 데이터 웨어하우스는 관련된 모든 사용자에게 데이터 요소가 수정되었다는 정보를 알려주어야 하기 때문에 데이터 요소에 대한 일련의 공정 및 갱신 결과를 추적하고 관리하는 것이 필요하다.

또한 시맨틱 웹(Semantic Web)은 컴퓨터가 정보를 구성하는 자원의 뜻을 이해하고 추론할 수 있는 지능형 웹을 의미하는 데, 이를 구현하기 위해서는 메타데이터를 통해 정보의 의미를 이해하고 처리하도록 하는 자원 설명 기술과 지식 설명 기술을 결합하는 과정이 필요하다. 정보통신 및 정보처리 분야에서 메타데이터 표준은 ISO(International Organizations for Standardization), ISO/IEC(International Engineering Consortium), W3C(The World Wide Web Consortium) 등과 같은 국제 표준화 기구에서 주도하고 있으며, 여러 정보통신 및 정보처리 분야에서 다양한 기술요소 및 어플리케이션 분야에 활용될 수 있는 메타데이터 표준을 개발하고 있다(Fig. 1).

우리나라에서도 국가 GIS 기술 개발 사업의 추진 초기부터 이러한 공간정보 메타데이터의 중요성을 인식하여 한국표준협회(KSA: Korea Standard Association)를 중심으로 ISO 국제 표준을 수용하거나, 한국정보통신기술협회인 TTA (Telecommunications Technology Association) 등을 중심으로 메타데이터 표준을 우리나라 환경에 맞게 개발하는 등과 같이 다양하고 적극적인

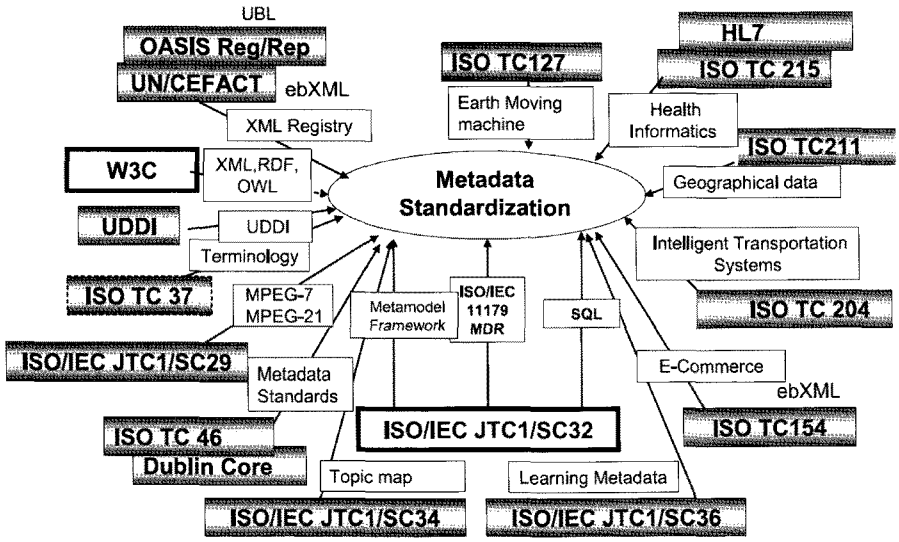

Fig. 1. Types of standard metadata in the information and communication domains in Seo (2008).

작업이 진행된 바 있다.

메타데이터와 관련된 기존의 연구로는 김 계현 등(2003), 이상문과 서정민(2007), 김 경민 등(2008)과 같이 표준 모델을 정리하거나 실제 활용된 사례 연구가 있다. 또한 이기원과 강혜경(2010)은 공간영상정보와 관련된 국내외 표준을 정리하면서 공간영상정보 메타데이터 표준 현황을 제시한 바 있다. 한편 Moellering *et al.*(2005)은 십여 개국의 국가 주도 메타데이터 구축 현황과 함께 국가별 공간정보 메타데이터 표준을 비교, 정리하였다. 그러나 그 동안 국내외에서 메타데이터와 관련하여 발표된 대부분의 연구는 이와 같이 표준 개발이나 활용 모델을 주요한 주제로 다루는 경우가 많았고, 본 연구의 주제 중의 하나인 메타데이터 추출에 관한 연구는 2000년대 중반부터 여러 사례가 발표되고 있다.

Greenburg (2004)는 공간정보를 대상으로 하지는 않았지만 일반적인 관점에서 메타데이터 구축 과정에 적용되는 추출과 수집(Harvest) 방법을 비교하면서 각각의 유용성을 비교한 바 있다. Diaz *et al.* (2007)은 영상 데이터와 출력된 지도에서 직접 데이터 형상과 관련된 메타데이터를 추출하는 연구를 수행한 바 있다. 한편 메타데이터의 중요성이 강조되면서 Wei *et al.*(2007)은 서로 다른 공간자료로부터 메타데이터 요소를 상호 변환하는 프로그램 개발에 대한 연구를 발표한 바 있다. 또한 Batcheller (2008)는 데이터 정보와 기존에 구축된 메타데이터를 연계하여 메타데이터를 갱신하고 추출하는 연구를, Schindler and Diepenbroek(2008)는 XML 기반의 메타데이터 포털 구축에 관한 프레임워크를 제시한 바 있다. Christel (2009)은 멀티미디어 동영상 정보에서 메타데이터의 자동 추출 방법을 정리하고 다양한 활용 사례를 발표한 바 있다. Kalantari *et al.* (2009)은 공간자료 데이터 포맷을 GML로 변환한 뒤에 이로부터 XML 형식의 메타데이터를 추출하는 연구를 수행한 바 있다.

그러나 기존의 연구 사례에서 보듯이 실무적인 관점에서 직접 공간자료로부터 메타데이터를 자동으로 추출하거나 메타데이터 표준간의 변환을 위한 연구는 국내외를 막론하고 그 실무적인 필요성에 비하여 연구된 사례가 많지 않다. 이는 현재까지도 메타데이터 구축은 실무 담당자가 직접 입력하는 경우가 일반적이며, 또한 메타데이터 사양도 국제 표준이나 국가, 단체 표준을 그대로 수용하기 보다는 구축 목적에 부합되는 최소의 항목만 다루거나, 운영 환경에 맞게 필수 항목만을 중심으로 설정하는 경우가 많기 때문이다. 그러므로 기존의 방법으로 구축 및 입력된 메타데이터는 입력된 항목을 이해하는 데 전문적인 지식이 필요하거나 입력시 의도하지 않은 오기나 오류가 발생할 가능성이 있다. 일부 상용 프로그램에서도 입력 데이터로 부터 좌표계, 데이터의 지리적 범위, 공간해상도 등과 같은 기본적인 메타데이터 요소가 자동 추출되는 기능이 있으나, 상용 프로그램에서는 국제 표준이나 단체 표준에 포함된 메타데이터

항목을 직접 지원하지 않고 별도의 사용자 입력을 요구하는 경우가 일반적이다.

본 연구에서는 벡터형 공간정보에 대해서는 국내 단체 표준인 TTA의 두 가지 메타데이터 표준과 ISO 국제 표준을 대상으로 하고, 공간영상정보의 경우에는 국제 표준을 대상으로 하여 공간정보 데이터로부터 직접 메타데이터를 자동으로 추출하고, 추출된 메타데이터를 서로 다른 표준으로 변환하는 프로그램 개발에 관한 연구를 수행하였다. 공간영상정보의 경우 한국정보통신기술협회(2005)에서 개발한 단체 표준인 '그리드 데이터 유통 목록(메타데이터)'이 있으나 이 단체 표준은 ISO/TC 19115를 기본 모델로 하여 확장한 것이므로 본 연구에서는 일단 고려 대상에 포함하지 않고 국제 표준만으로 적용하였다. 물론 공간영상정보를 주로 다루는 응용 시스템에서는 국제 표준인 ISO/TC 19115-2와 단체 표준인 TTAS.KO-10.0194를 모두 처리해야 한다. 또한 연구 결과로 설계 구현한 프로그램의 구성 요소와 활용 절차를 제시하고자 한다.

2. 메타데이터 사양 비교

공간정보 메타데이터는 실무적인 수요에 따라 표준 개발을 위한 연구와 표준화 작업이 1990년대 중반부터 수행되었고, ISO/TC 211에서 메타데이터 표준을 개발하여 공표한 이후 가장 활용도가 높은 표준으로 알려져 있다.

Moellering *et al.*(2005)은 ISO/TC 19115 표준을 기준으로 하여 국가별 공간정보 메타데이터 표준을 객관적으로 비교한 바 있다. 대부분의 국가에서는 공간정보 메타데이터 표준을 ISO/TC 19115에 기반하고 있으나 국가별 상황에 맞게 다양한 프로파일이 개발, 적용되고 있다. 우리나라의 경우도 마찬가지 상황으로, Table 1, Table 2, Table 3에서 몇 가지 내용을 비교하고자 하였다. 그러나 ISO/TC 19115 표준이 주로 벡터 자료를 대상으로 하므로 비교 과정에서 공간영상정보와 격자형 자료에 관련된 국제 표준인 ISO/TC 19115-2 표준 내용은 포함하지 않았다. 또한 여기서 ISO/TC 19115는 영문 그대로 기술하였으며, TTA의 유통 및 관리용 표준은 각

Table 1. Comparison of standard metadata: Metadata entity sets

	ISO/TC 19115	TTAS.KO-10.0139(유통)	TTAS.IS-19115(관리)
요소 (Element)	fileIdentifier	메타데이터 파일식별자	메타데이터 파일 식별자
	language	메타데이터 언어	메타데이터 언어
	characterSet	메타데이터 문자셋	메타데이터 문자셋
	hierarchyLevel	메타데이터 적용계층 대상	메타데이터 적용계층 대상
	hierarchyLevelName	메타데이터 적용계층 대상명	메타데이터 적용계층 대상명
	contact	메타데이터 연락정보	메타데이터 연락정보
	dateStamp	메타데이터 생성일자	메타데이터 생성일자
	metadataStandardName	메타데이터 표준명	메타데이터 표준명
	metadataStandardVersion	메타데이터 표준버전	메타데이터 표준버전
	parentIdentifier	(없음)	(없음)
개체 (Entity)	MD_ReferenceSystem	MD_참조체계	MD_참조체계
	MD_MaintenanceInformation	MD_유지관리	MD_유지보수
	MD_Identification	MD_식별	MD_식별
	MD_Distribution	MD_배포	MD_배포
	DQ_DataQuality	DQ_데이터품질	DQ_데이터품질
	MD_SpatialRepresentation	(없음)	MD_공간표현
	MD_MetadataExtensionInformation	(없음)	MD_메타데이터확장정보
	MD_Constraints	(없음)	MD_제약
	MD_ApplicationSchemaInformation	(없음)	MD_응용스키마정보
	MD_PortrayalCatalogueReference	(없음)	(없음)
	MD_ContentInformation	(없음)	MD_내용정보

각 구분하였다.

Table 1은 메타데이터 표준 내용 중 메타데이터 개체 셋에 대한 메타데이터 항목을 비교한 것이다. 관리용 공간정보 메타데이터 표준인 TTAS.IS-19115는 ISO/TC 19115를 기본 모델로 하고 있으므로 parentIdentifier나 PortrayalCatalogueReference 등과 같은 일부 요소를 제외한 대부분의 항목은 일치하고 있다. 그러나 TTA 유통 메타데이터 표준의 경우에는 TTA 관리용 표준 항목에 포함되어 있는 '공간표현', '제약', '내용정보' 등이 제외되어 있음을 알 수 있다.

Table 2는 메타데이터 표준 내용 중 식별 정보에 대한 항목을 나타낸 것으로 ISO/TC 19115와 TTA 관리용 메타데이터는 geographicBox나 geographicDescription 등을 제외하고 많은 부분이 대응하고 있다. 그러나 ISO/TC 19115와 TTA 유통용 메타데이터는 이러한 항목이 '지리경계', '지리설명' 등으로 번역되어 대응하고 있다. 그러나 TTA 유통용 메타데이터는 개체 요소에서 '주제어', '용도', '제약' 등의 항목은 제외되어 있음을 알 수 있다.

Table 3은 메타데이터 표준 내용 중 식별 정보가 참조하는 개체 셋에 대한 내용이다. ISO/TC 19115를 기준으로 하여 TTA 관리용 메타데이터는 '파일명', '주제어', '거리', '상세용도', '사용자연락정보' 등이 번역되어 대응하고 있지만 TTA 유통용 메타데이터에는 해당 항목들이 제외되어 있음을 알 수 있다. 또한 ISO/TC 19115의 denominator, equivalentScale은 TTA 유통

Table 2. Comparison of standard metadata: Identification

	ISO/TC 19115	TTAS.KO-10.0139(유통)	TTAS.IS-19115(관리)
요소 (Element)	MD_Identification(MD_식별)		
	citation	참고자료	참고자료
	abstract	요약설명	요약설명
	purpose	(없음)	목적
	credit	(없음)	(없음)
	status	(없음)	상태코드
	postOfContact	공간정보연락처	공간정보연락처
	MD_DataIdentification(MD_데이터식별)		
	spatialRepresentationType	공간표현유형	공간표현방식
	spatialResolution	공간해상도	공간해상도
	language	자원언어	자원언어
	characterSet	자원문자셋	자원문자셋
	topicCategory	주제분류	주제분류
	geographicBox	지리경계	(없음)
	geographicDescription	지리설명	(없음)
	environmentDescription	(없음)	환경설명
	extent	범위	범위
	supplementalInformation	(없음)	보충정보
	(없음)	검색정보	검색정보
개체 (Entity)	MD_Identification(MD_식별)		
	MD_MaintenanceInformation	(없음)	MD_유지보수정보
	MD_BrowseGraphic	(없음)	MD_도시용그래픽
	MD_Format	MD_포맷	MD_포맷
	MD_Keywords	(없음)	MD_주제어
	MD_Usage	(없음)	MD_용도
	MD_Constraints	(없음)	MD_제약
	MD_DataIdentification(MD_데이터식별)		
	(없음)	DT_자료구조형태	DT_자료구조

Table 3. Comparison of standard metadata: Reference entity sets of identification

	ISO/TC 19115	TTAS.KO-10.0139(유통)	TTAS.IS-19115(관리)
요소 (Element)	MD_BrowseGraphic(MD_도시용그래픽)		
	fileName	(없음)	파일명
	fileDescription	(없음)	(없음)
	fileType		(없음)
	MD_Keywords(MD_주제어)		
	keyword	(없음)	주제어
	type	(없음)	(없음)
	thesaurusName	(없음)	(없음)
	MD_RepresentativeFraction(MD_대표비율)		
	denominator	분모	비율
	MD_Resolution(MD_해상도)		
	equivalentScale	상응스케일	동치축척
	distance	(없음)	거리
	MD_Usage(MD_용도)		
	specificUsage	(없음)	상세용도
	usageDateTime	(없음)	(없음)
	userDeterminedLimitations	(없음)	(없음)
	userContactInfo	(없음)	사용자연락정보
	DT_DataStructure(DT_자료구조)		
	DT_Database(DT_데이터베이스)		
	(없음)	데이터제공자종류	데이터제공자종류
	(없음)	주제명	주제명
	(없음)	레이어명	레이어명
	DT_File(DT_파일)		
	(없음)	파일경로	파일경로
	DT_DistributionInformation(DT_유통정보)		
	(없음)	도엽명	도엽명
	(없음)	도엽번호	도엽번호
	(없음)	지도분류유형	지도분류유형
	(없음)	지역범위	지역범위

및 관리 메타데이터에서 각각 '분모'와 '비율', '상응스케일'과 '동치축척'으로 다르게 번역되어 대응하고 있다. 한편 '데이터베이스'와 '유통정보'와 같은 개체 셋은 TTA 메타데이터에 프로파일링되어 추가된 개체이므로 국제 표준인 ISO/TC 19115에는 없음을 알 수 있다.

물론 이 세 가지 비교 항목은 메타데이터의 전체 중 일부에 해당하며, 다른 항목에서도 상호 간에 대응하지 않거나 일치하지 않는 경우가 많이 있으나, TTA의 메타데이터 표준 문서에 배경이나 근거가 제시되어 있으므로 본 연구에서는 표준항목만을 고려하고자 하였다.

한편 Fig. 2는 공간 데이터 포맷을 비교한 Manso *et al.*(2004)의 자료를 인용한 것으로 벡터형 자료와 공간영상정보 포맷을 포함하는 격자형 자료를 구분하여 나타낸 것이다. 격자형 자료 포맷 내에 포함되어 있는 BoundingBox, Pixel Resolution, Projection, Datum, Ellipsoid 등과 같은 요소는 메타데이터 표준 항목에도 거의 포함되어 있는 내용이므로 사용자가 별도로 입력하지 않아도 프로그램에서 바로 읽어서 메타데이터로 자동 입력이 가능한 요소가 될 수 있다. 물론 이 자료가 세부적인 내용을 포함하지 않고 파일 포맷 자체도 계속 갱신되거나 추가되고 있으므로 보다 자세한 비교가 필요하다.

Types of Geo-spatial Data Set in Vector Model

Format	BndBox	N° of layers	Name of layers	N° of different features	Name & number of features	Horizontal Units	Projection	Date	Ellipsoid
E00 arc	X								
ADF arc	X			X					
DGN	X	X		X					
DXF	X	X	X	X					
SHP	X	X	X	X					
MIF	X			X					
TAB	X			X					
DWG	X	X	X	X					
VEC (idrisi)	X			X					
BIN	X	X[5]	X[5]	X					

Types of Geo-spatial Image or Gridded Data Set

Format	W / H	BoundBox	pixel resolution	Bits / pixel	Bands / dimensions	Max, min statistical	Other stadistics	Horizontal Units	Projection	Datum	Ellipsoid	Other
Gif	X	X[1]	X[1]	X	X							
Png	X	X[2]	X[2]	X	X							
Jpg	X	X[1]	X[1]	X	X							
Iff	X	X[1]	X[1]	X	X							
Tiff	X	X[1]	X[1]	X	X							
GeoTiff	X	X	X	X	X			X	X	X	X	
Bmp	X	X[1]	X[1]	X	X							
Pcx	X	X[1]	X[1]	X	X							
Psd	X	X[3]	X[3]	X	X							
Ras	X	X[1]	X[1]	X	X							
Sid	X	X[1][2]	X[1][2]	X	X			X[2]	X[2]			
Ecw	X	X[1][2]	X[1][2]	X	X			X[2]	X[2]	X[2]		
JP2000	X	X[1]	X[1]	X	X							
GeoJP2(JP2k)	X	X	X	X	X			X	X	X	X	
DOQ2 (USGS)	X	X	X	X	X			X	X	X		
DTED	X	X	X	X	X			X	X	X	X	
DTE-Socet set	X	X	X	X	X			X				
GRD ESRI	X	X	X			X[2]		X	X	X	X	
GRD surfer	X	X	X			X						
DEM -USGS	X	X	X	X		X		X	X	X		
MicroDEM	X	X	X	X	X	X		X	X	X		
E00 grd	X	X	X[4]		X	X[4]	X[4]	X[4]	X[4]	X[4]	X[4]	
NTIF	X	X	X	X	X			X	X			X
ADF grd	X	X	X[4]		X	X[4]						
Lan erdas	X	X	X	X	X							
IMG erdas	X	X	X	X	X	X						
PIX	X	X	X		X							
ERS	X	X	X	X	X	X						
DOC (Idrisi)	X	X	X	X		X		X	X	X	X	

Fig. 2. Metadata elements within Geo-based file formats, excerpted from Manso *et al.*(2004).

```
Geotiff_Information:
   Version: 1
   Key_Revision: 1.0
   Tagged_Information:
      ModelTiepointTag (2,3):
         0                 0
         742295.721        4278516.24
      ModelPixelScaleTag (1,3):
         1                 1
      End_Of_Tags.
   Keyed_Information:
      GTModelTypeGeoKey (Short,1): M
      GTRasterTypeGeoKey (Short,1):
      GeographicTypeGeoKey (Short,1)
      GeogAngularUnitsGeoKey (Short,
      ProjectedCSTypeGeoKey (Short,1
      ProjLinearUnitsGeoKey (Short,1
      End_Of_Keys.
   End_Of_Geotiff.

PCS = 32615 (name unknown)
Projection = 16015 ()
Projection Method: CT_TransverseMerc
   ProjNatOriginLatGeoKey: 0.000000
   ProjNatOriginLongGeoKey: -93.0000
   ProjScaleAtNatOriginGeoKey: 0.999
   ProjFalseEastingGeoKey: 500000.00
   ProjFalseNorthingGeoKey: 0.000000
GCS: 4326/WGS 84
Datum: 6326/World Geodetic System 19
Ellipsoid: 7030/WGS 84 (6378137.00,6
Prime Meridian: 8901/Greenwich (0.00
Projection Linear Units: 9001/metre

Corner Coordinates:
Upper Left    (  742295.721, 4278516
Lower Left    (  742295.721, 4274733
Upper Right   (  745950.721, 4278516
Lower Right   (  745950.721, 4274733
Center        (  744123.221, 4276624
```

```
INST_LAST_MUC_DATE      20061216
INST_LAST_GEO_DATE      20080220
INST_COMPRESSION_FLAG   TRUE
INST_COMPRESSION_RATIO_OF_MS    0 0 0 0
INST_TDI_GAIN_OF_MS     3 4 1 2
INST_ELEC_GAIN_OF_MS    0 0 0 0
INST_ELEC_OFFSET_OF_MS  0 0 0 0
INST_BAND_DISPLAY       G
INST_BAND_WIDTH 80
INST_MS_CCD_ALIGNMENT     -0.0247100      -0.0226570      0.
INST_MS_FOCAL_LENGTH      2.25550000
INST_CCD_MODE   MS_Primary
CAL_MTF_OF_MS   -1047013500 0 -1047013500 0
CAL_RADIANCE_GAINOFFSET_MS        1.00   1.00   1.00   1.00
BEGIN_CALGCP_BLOCK
        NMR_GCP 0
END_CALGCP_BLOCK
CAL_DEM_FILE    NULL
AUX_FILE_NAME   MSC_09040201290Z_14297_10801272M1N04G_1G.tif
AUX_STRIP_ID    MSC_09040201290Z_14297_10811255_1404BN04_1A
AUX_STRIP_BEGIN_END     61968 65842
AUX_IMAGE_LEVEL L1G
AUX_PRODUCT_LEVEL       RPC
AUX_CLOUD_STATUS        0
AUX_IMAGE_QUALITY       NULL
AUX_IMAGE_BAD_LINES     NULL
AUX_IMAGE_BAD_COLS      NULL
AUX_SATELLITE_NAME      KOMPSAT2
AUX_SATELLITE_SENSOR    MSC
AUX_TILT_ANGLE_ROLL_DEG  -4.473
AUX_TILT_ANGLE_PITCH_DEG        -0.570
AUX_BITS_PER_PIXEL      16
AUX_SAMPLES_PER_LINE_MS 4445
AUX_LINES_PER_IMAGE_MS  4571
AUX_SCENE_CENTER_XY_PIXEL       2222 2285
AUX_GSD_METER   4.047 3.906
AUX_IMAGE_GSD_METER     4.000 4.000
AUX_LINE_SCAN_TIME_USEC 0.0004927437
AUX_IMAGE_SATELLITE_AZIMUTH_DEG 93.426072
AUX_IMAGE_SATELLITE_INCIDENCE_DEG       5.188612
AUX_IMAGE_PAD_POD_FLAG  TRUE
AUX_IMAGE_MTF_FLAG      FALSE
```

Fig. 3. Metadata elements within Geo-Tiff file: Example of KOMPSAT-2 image sets.

Fig. 3은 KOMPSAT-2 영상 자료에 포함되거나 별도로 제공되는 메타데이터 정보를 나타낸 것이다. 영상 자료 자체를 이용하는 활용 연구에서는 일부 항목만 필요로 할 수 있으나, 공간영상정보를 대상으로 하는 공간데이터베이스를 구축하거나 이와 관계된 검색 시스템이 필요한 경우에는 여기서 제시된 거의 모든 항목에 대한 내용을 메타데이터 요소로 입력하고 저장, 관리할 필요가 있다. 따라서 이러한 내용도 자료 내에 포함되어 있거나 일반 파일 형식으로 제공되고 있으므로 사용자가 입력하기 보다는 자동으로 처리하는 것이 효율적임을 알 수 있다. 또한 이러한 메타데이터 정보는 데이터 제공자가 사용자의 주관적 판단을 요구하지 않도록 별도의 설명 및 안내 자료를 제공하고 있고, 대부분의 요소가 이미 공간정보 메타데이터 표준 항목으로 설계되어 있으므로 자동 처리가 충분히 가능한 내용이라고 할 수 있다.

3. 프로그램 구현

본 연구의 프로그램 개발 구성 또는 활용 절차는 Fig. 4와 같이 정리할 수 있다. 본 연구에서 구현한 프로그램 모듈은 공간자료 포맷 및 별도로 제공되는 메타데이터 정보를 읽는 Geo-based Data Sets Importer, 공간자료 포맷 및 별도로 제공되는 메타데이터 정보에서 자동으로 메타데이터 표준 형식으로 항목을 추출하도록 하는 Metadata Extractor, 추출된 메타데이터 항목을 다시 XML 형식으로 표현하는 Metadata XML Exporter, 사용자가 요구하는 메타데이터 표준 간의 상호 변환을 수행하는 Metadata Transformer, 자동 입력된 메타데이터를 확인하고 누락된 내용이나 변경되는 사항을 사용자가 수정, 갱신하도록 하는 Metadata Editor 등으로 이루어져 있다. 활용 절차의 경우는 ①, ②, ③, ④, ⑤와 같은 순서로 Fig. 4에 제시하였다.

프로그램 개발환경은 Visual Studio C++이며, XML 파싱 및 생성 시에는 TinyXML 라이브러리를 사용하였으며, 공간자료 포맷으로부터 메타데이터를 추출하는 처리에서는 GeoTiff_Lib와 ShapeLib등과 같은 오픈 소스를 사용하였다.

Fig. 5는 Metadata Extractor에 포함된 사용자 환

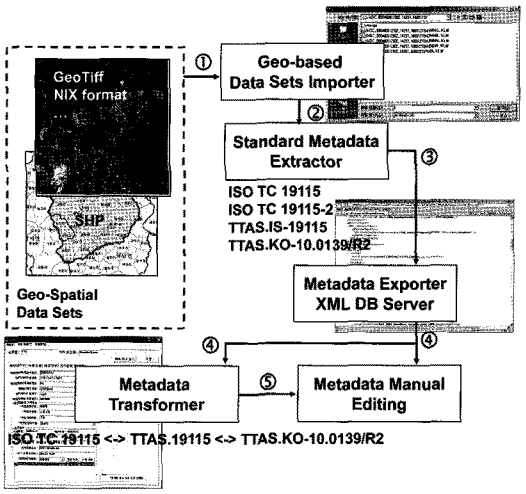

Fig. 4. Work processes for metadata extraction and transformation using program implemented in this study.

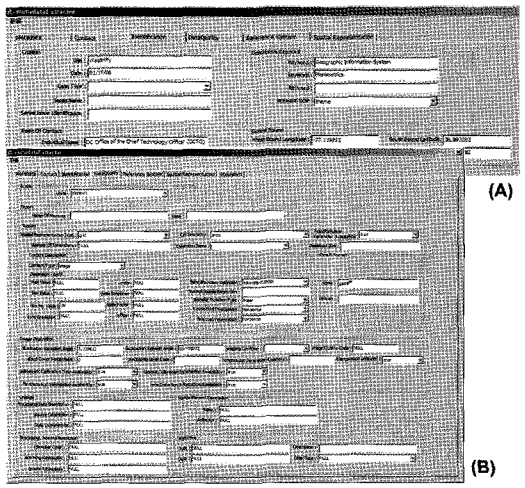

Fig. 5. User interfaces for metadata extraction: (A) Shapefile for vector format, (B) Geo-Tiff for geo-image format.

경의 하나로, (A)는 벡터 형 공간자료 포맷인 Shapefile을 읽어 ISO/TC 19115 형식으로 메타데이터에 적용한 사례이고, (B)는 Geo-Tiff 파일을 읽어서 이로부터 ISO/TC 19115-2 형식으로 메타데이터를 추출한 사례이다. 물론 프로그램 구동 시에는 사용자가 별도로 메타데이터 표준을 지정하지 않아도 TTA의 관리용, 유통용 메타데이터 항목도 자동으로 추출하도록 하였다. 이는 사용자가 추후에 메타데이터 표준간의 변환 처리를 가능하도록 한 것으로, Table 1, Table 2, Table 3에서 설명한 바와 같이 메타데이터 표준 간에도 서로 일치하지 않거나 대응하지 않는 요소들이 존재하기 때문이다. 또

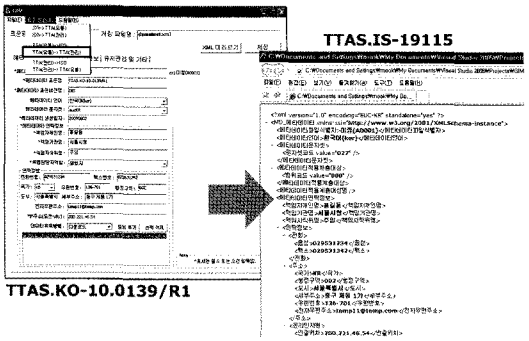

Fig. 6. Metadata transformation: TTA.KO-10.0139/R1 to TTAS.IS-19115.

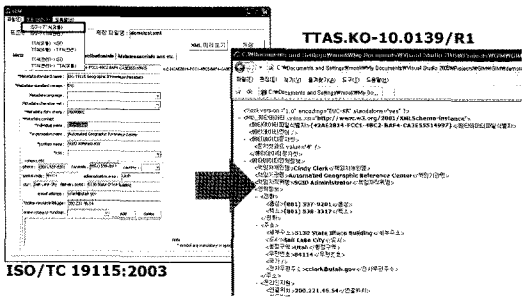

Fig. 7. Metadata transformation: ISO/TC 19115 to TTA.KO-10.0139/R1.

한 공간자료로부터 직접 메타데이터를 추출하는 경우에도 표준 항목에 대응하는 정보가 없거나 일치하지 않는 경우가 발생하는 데, 이와 같은 경우에는 NULL 값으로 처리하였고 담당자가 추출 결과의 검증 단계에서 이를 수정하거나 새로 입력할 수 있도록 하였다.

Fig. 6과 Fig. 7은 각각 TTA 유통용 메타데이터 표준을 관리용 메타데이터로 변환하거나 ISO/TC 19115 메타데이터를 TTA 유통용 메타데이터 표준으로 변환하도록 하는 사용자 환경의 일부를 제시한 것이다. 이러한 경우에도 최초 추출된 메타데이터에서 NULL 값으로 입력된 사항을 담당자가 별도로 수정하지 않은 경우에는 그대로 유지가 되나, 원래 공간자료에 있던 정보가 메타데이터 항목에 없어 입력되지 못하였으나 다른 메타데이터 표준에서 이에 해당하는 항목이 설정되어 있는 경우에는 변환 후에는 입력이 되도록 하였다.

그러나 벡터 유형의 공간자료를 다루는 경우와 공간영상자료를 다루는 경우에는 각각의 메타데이터 표준항목이 필수 기본 요소를 제외하고는 상이하기 때문에 프로그램 환경에서는 벡터 자료와 공간영상자료 간의 변환을 위한 사용자 환경은 개발 내역에 별도로 포함하지 않았다.

한편 ISO/TC 19115는 영어로 주로 입력이 되고, TTA 표준은 국문 단어로 처리되어야 하기 때문에 Fig. 7과 같이 XML 요소는 국문으로, 입력 결과는 영어로 처리 결과가 나타나게 된다. 이는 국영문 변환 처리 용어 사전(Data Dictionary)을 별도로 구축하거나 단어 번역기 등과 같은 추가 기능을 통하여 해결될 수 있으며, 온톨로지 기반 메타데이터 추출 및 변환 시스템 개발이 실무적으로도 필요하다는 것을 나타내는 것이다.

4. 결 론

일반적으로 공간정보 메타데이터 표준은 일관성 있고 체계적인 정보의 제공, 실 데이터 선택의 편의성 제공, 정보 자원의 호환성과 상호 운용성을 증진하는 등과 같은 다양한 목적을 위한 핵심 요소이다. 또한 웹 기반 컴퓨팅 기술과 정보통신 운영 환경이 발전하면서 그 중요성이 더욱 증가되고 있고, 다양한 메타데이터 표준이 개발되고 공표되고 있다. 그러나 공간정보 메타데이터 표준은 생성규칙의 어려움, 소스 및 표준의 다양성, 동기화된 데이터의 유지 등과 같은 운영상의 문제점이 있어 실무적인 활용에 제약이 되기도 한다. 또한 메타데이터 서버를 구축하는 경우에는 XML 형식으로 저장, 관리할 필요가 있기 때문에 실무 응용 시스템에서는 표준 형식을 따르기 보다는 운영 기관별로 별도의 메타데이터 사양을 설정하여 사용하는 경우가 많다. 본 연구에서는 국제 표준인 ISO/TC 19115와 TTAS.KO-10.0139와 TTAS.IS-19115와 같은 단체 메타데이터 표준을 손쉽게 입력할 수 있도록 직접 공간자료 포맷으로부터 메타데이터 항목과 요소를 추출하여 직접 XML로 저장할 수 있는 프로그램을 개발하였다. 이 세 가지 메타데이터 표준은 ISO/TC 19115에서 정의한 22개의 필수 핵심 항목은 동일하지만 프로파일이 상이하여 많은 부분이 일치하지 않거나 대응하지 않는다. 즉, 메타데이터 표준 간의 항목이나 요소들이 일부 상이하기 때문에 메타데이터 추출 기능과 함께 메타데이터 간의 변환 기능을 개발하여 메타데이터의 상호 운영이 가능하도록 하였다. 또한 공간영상정보에 대한 자료 포맷의 경우에는 ISO/TC 19115-2 메타데이터 표준 사양으로 추출, 저장할 수 있

도록 하였다. 특히 메타데이터 자동 추출의 경우 메타데이터 표준의 100 여 가지가 넘는 다양한 항목(필수, 선택, 확장) 중 자동 추출의 대상이 되는 항목은 공간 데이터 자체에서 포함된 내용만을 우선적인 추출 대상으로 인식한다. 따라서 메타데이터 표준 항목 중 임의 서술을 허용하고 있는 항목은 사용자 입력 방식으로 처리될 수 밖에 없다. 그러나 최근 공간정보 데이터 생산이나 제작 과정이나 공간정보 자료 포맷 자체에서 표준 항목과 부합되는 점차 메타데이터 항목이 포함되는 경우가 많아지고 있기 때문에 공간정보 응용 시스템 설계나 구축 단계에서 메타데이터 자동 추출 기능에 대한 필요성도 증대할 것으로 생각한다. 또한 점차 그 수요와 필요성이 증가하고 있는 여러 가지 기존에 구축된 응용 시스템 간의 연계와 통합 과정에서 국제 표준과 단체 표준 간의 변환 기능도 유용하게 적용될 수 있다.

본 연구에서는 활용도가 높은 자료 포맷을 대상으로 하였으나 기타 자료 포맷이나 기관별 메타데이터 사양을 처리할 수 있는 확장 기능 구현을 통하여 메타데이터 표준의 활용도를 증가시킬 수 있을 것으로 기대한다.

사 사

본 연구는 국토해양부 첨단도시기술개발사업 – 지능형국토정보기술혁신 사업과제의 연구비지원(과제번호: 07국토정보C03)에 의해 수행되었습니다.

참고문헌

김경민, 김철민, 김태균, 2008. 수치임상도 표준 메타데이터 설계 및 구현, 한국지리정보학회지, 11(4): 51-63.

김계현, 송용철, 김한국, 민숙주, 2003. 지리정보의 활용을 지원하기 위한 메타데이터 표준안 정립에 관한 연구, 개방형지리정보시스템학회지, 5(2): 55-68.

박용재, 이기원, 2010. 공개소스 DBMS 미들웨어 연동 공간정보 브라우저 설계 및 프로토타입 구현, 대한원격탐사학회지, 26(2): 99-108.

이기원, 강혜경, 2010. 공간영상정보 관련 ISO와 OGC 표준현황과 활용을 위한 제언, 대한원격탐사학회지, 26(4): 451-464.

이상문, 서정민, 2007. 지하수관리시스템의 공간 메타데이터 모델에 관한 연구, 한국컴퓨터정보학회 논문집, 12: 229-237.

서 태설, 2008. 메타데이터, TTA Journal, 119, 113-118.

한국정보통신기술협회, 2005. 그리드데이터 유통목록(메타데이터), TTAS.KO-10.0194.

한선묵, 이기원, 2010. 지형공간정보 카탈로그 서비스 기본요소의 PostgreSQL 연동 시험모델 구현, 대한원격탐사학회지, 26(2): 133-142.

Baca, M., 2008. *Introduction to Metadata (2nd)*, Getty.

Batcheller, J. K., 2008. Automating geospatial metadata generation?An integrated data management and documentation approach, *Computers & Geosciences*, 34: 387-398.

Christel, M. G., 2009. *Automated Metadata in Multimedia Information Systems: Creation, Refinement, Use in Surrogates, and Evaluation*, Morgan & Claypool, 73p.

Diaz, L., C. Martin, M. Gould, M. A. Manso, 2007. Semi-automatic Metadata Extraction from Imagery and Cartographic data, *IGARSS 2007*.

Kalantari, M., A. Rajabifard and H. Olfat, 2009. Spatial metadata automation : a new approach, *Proceedings of the Surveying & Spatial Sciences Institute Biennial International Conference*: 629-635.

Lawrence A., W., Traci, and J. Hess, 2002. Metadata as a knowledge management tool: supporting intelligent agent and end user access to spatial data, *Decision Support Systems*, 32: 247-264.

Manso, M. A., J. Nogueras-Iso, M. A. Bernabe, F. J. Zarazaga-Soria, 2004. Automatic Metadata Extraction from Geographic Information, *7th AGILE Conference on Geographic Information*

Science.

Moellering, H., H. J. G. Aalders, and A. Crane (eds), 2005. *World Spatial Metadata Standards: Scientific and Technical Descriptions, and Full Descriptions with Crosstable*, Elsevier, 689p.

Greenberg, J., 2004. Metadata Extraction and Harvesting: A Comparison of Two Automatic Metadata Generation Applications, *Journal of Internet Cataloging*, 6(4): 59-82.

Schindler, U. and M. Diepenbroek, 2008. Generic XML-based framework for metadata portals, *Computers & Geosciences*, 34: 1947-1955.

Wei, Y., L. Di, B, Zhao, G. Liao, and A. Chen, 2007. Transformation of HDF-EOS metadata from the ECS model to ISO 19115-based XML, *Computers & Geosciences*, 33: 238-247.