

임산공학 분야 연구데이터 관리를 위한 메타데이터 설계에 관한 연구*

A Study on the Design of Metadata for Research Data Management in Forestry Engineering

김 주 섭 (Juseop Kim)**, 한 연 중 (Yeonjung Han)***
유 원 재 (Wonjae Youe)****, 김 선 태 (Suntae Kim)*****

목 차

- | | |
|------------------------------|--|
| 1. 서 론 | 4. 임산공학 분야 연구데이터 관리를 위한
메타데이터 요소 선정 |
| 2. 이론적 배경 | 5. 결 론 |
| 3. 임산공학 분야 관련 메타데이터
요건 분석 | |

초 록

최근 많은 학문 분야에서 해당 분야에서 생산되는 데이터를 관리하기 위한 메타데이터를 개발하고 있다. 범용적인 분야를 넘어서 특정 학문 분야를 지원하기 위한 메타데이터는 해당 분야의 연구데이터를 검색, 관리, 공유 및 재사용을 위한 필수한 것으로 여겨지고 있다. 따라서 본 연구에서는 임산공학에서 생산되는 연구데이터를 관리하기 위한 메타데이터 항목을 도출하였다. 메타데이터 항목을 도출하기 위하여 FGDC-CSDGM, ISO 19115, NEFIS, INSPIRE, ANZLIC, DataCite 4.3 그리고 TTA.KO-10.0976 등 7개 메타데이터를 대상으로 매핑작업을 진행하였다. 연구 결과, 필수 6개, 권고 13개 그리고 선택 요소 9개가 도출되었다. 도출된 메타데이터는 임산공학 분야의 연구데이터를 체계적으로 관리하기 위한 항목으로 활용될 수 있을 것이다.

ABSTRACT

Recently, many academic fields are developing metadata to manage data produced in the field. Metadata to support a specific academic field beyond the general field is considered essential for searching, managing, sharing, and reusing research data in the field. Therefore, in this study, metadata items were derived to manage research data produced in forestry engineering. In order to derive metadata items, mapping work was performed on 7 metadata such as FGDC-CSDGM, ISO 19115, NEFIS, INSPIRE, ANZLIC, DataCite 4.3, and TTA.KO-10.0976. As a result of this study, 6 mandatory, 13 recommendations and 9 optional elements were derived. The derived metadata can be used as an item to systematically manage research data in the field of forestry engineering.

키워드: 메타데이터, 연구데이터, 크로스워크, NEFIS, DataCite
Metadata, Research Data, Crosswalk, NEFIS, DataCite

- * 본 연구는 2019년도 국립산림과학원의 연구지원에 의해 이루어진 것임.
이 논문은 2020년도 전북대학교 연구기반 조성비 지원에 의하여 연구되었음.
- ** 전북대학교 문헌정보학과 강사(kimjuseop@jbnu.ac.kr / ISNI 0000 0004 7492 1806) (제1저자)
- *** 국립산림과학원 목재이용연구부 임업연구사(yeonjung@korea.kr / ISNI 0000 0004 5947 1958) (공동저자)
- **** 국립산림과학원 목재이용연구부 임업연구사(sngkgk@korea.kr / ISNI 0000 0004 8032 4581) (공동저자)
- ***** 전북대학교 문헌정보학과 조교수(kim.suntae@jbnu.ac.kr / ISNI 0000 0004 6492 6355) (교신저자)
- 논문접수일자: 2020년 10월 20일 최초심사일자: 2020년 11월 5일 게재확정일자: 2020년 11월 17일
한국문헌정보학회지, 54(4): 169-194, 2020. <http://dx.doi.org/10.4275/KSLIS.2020.54.4.169>

1. 서론

1.1 연구의 필요성 및 목적

작은 센서로부터 거대 실험 장치들이 원유라 불리는 데이터를 실시간으로 생산하고 있다. 학술 연구 영역에서는 이러한 데이터를 '연구 데이터(research data)'라 부르며, 공적 펀딩을 받아 수행된 연구 산출물에 대해 법적, 제도적, 기술적 장벽 없이 접근과 활용이 가능하도록 하자는 오픈 사이언스(Open Science) 운동은 주요 선진국에서 다양한 형태의 활동을 이끌어 내고 있다. 오픈 액세스(Open Access), 오픈 데이터(Open Data), 오픈 리포지토리(Open Repository) 등으로 구체화되는 오픈 사이언스 운동에서도 데이터가 그 핵심에 있다(김주섭, 김선태, 최상기 2019).

데이터의 중요성이 증대 될수록 데이터에 대한 공유·활용에 대한 요구가 다양한 분야에서 제기되고 있다. 예를 들어, 기후 데이터의 경우, 과거에는 기후 분야를 연구하는 과학자들이 주 사용자였으나 오늘날에는 물과 공공용지(public land), 건강, 해양 분야 등의 자원 관리자들이 그들의 의사결정과 연구를 위해 기후 데이터에 대한 접근을 요청하고 있으며, 정책 수립 분야에서도 기후 변화 대응 전략을 수립하기 위하여 데이터에 대한 접근을 필요로 하며, 기하급수적으로 발생하는 기후 데이터에 대한 물리적인 보존과 공유에 대한 이슈가 활발하게 제기되고 있다.

『목재의 지속가능한 이용에 관한 법률』(이하 목재이용법)의 시행으로 산림청은 목재제품 품질관리 등의 정책을 추진하고 있다. 특히, 국

립산림과학원은 목재연구를 통해 친환경 접착제 연구, 친환경 목재보존처리 연구, 나노셀룰로오스를 응용한 복합신소재 연구, 목재의 난연화(難燃化) 연구와 관련된 데이터를 생산하고 있다. 이에 따라 체계적이고 지속가능한 목재·목재제품 품질관리를 위해 연구데이터의 관리 및 서비스가 반드시 요구되는데 이에 본 연구에서는 해당 분야에서 생산되는 연구데이터를 위한 메타데이터 항목을 도출하고자 한다. 도출된 메타데이터를 통해 국립산림과학원에서 생산되는 연구데이터를 관리 및 재사용하기 위한 플랫폼 설계가 가능할 것이며 향후, DMP(Data Management Plan) 작성에 해당 연구 결과를 활용할 수 있는 가이드가 될 수 있을 것이다.

1.2 연구 방법 및 절차

본 연구의 목적을 달성하기 위해 사용된 연구 방법 및 절차는 다음과 같다. 다음 〈그림 1〉은 본 연구의 내용 및 절차를 도식화한 것이다.

첫 번째 단계에서는, 전반적인 연구의 방향과 범위를 설정하고 두 번째 임산공학의 메타데이터를 도출하기 위하여 관련 표준을 조사 및 분석하였다. 임산공학 분야 메타데이터를 설계한 사례가 NEFIS를 제외하고 전무하여 본 연구에서는 가장 관련 있는 지리정보를 기술하기 위한 메타데이터인 FGDC-CSDGM 등과 연구데이터를 기술하기 위한 메타데이터인 DataCite를 중심으로 연구를 진행하였다. 세 번째 단계에서는 임산 공학 분야 연구데이터 관리를 위한 메타데이터 항목을 도출하였다. 해당 단계에서는 2단계를 통해 분석한 관련 분

1단계	연구 방향 및 범위 설정을 위한 브레인 스토밍
2단계	임산 공학 분야 메타데이터 분석: FGDC-CSDGM 등
3단계	임산 공학 분야 연구데이터 관리를 위한 메타데이터 도출
4단계	연구자 설문조사를 통한 메타데이터 요소 검증
5단계	최종 메타데이터 요소 도출 및 스키마 작성

〈그림 1〉 연구 방법 및 절차

야 메타데이터 표준을 대상으로 매칭 테이블을 통해 메타데이터 항목을 도출하였으며 도출된 메타데이터 항목은 검증을 위하여 관련 전문가를 대상으로 설문조사를 실시하였다. 마지막 단계에서는 4단계를 통해 수행된 설문조사 결과를 바탕으로 메타데이터 항목의 우선순위를 정하고 각 항목에 대한 메타데이터 스키마를 작성하였다.

위와 같이 임산공학 분야를 위한 메타데이터 요소를 도출하기 위하여 FGDC-CSDGM 등 7개 메타데이터의 최상위요소와 핵심요소를 분석하였다. 각 메타데이터의 모든 요소를 이용하여 임산공학 분야의 메타데이터 요소를 도출해야 하지만 일부 메타데이터의 경우 411¹⁾개의 요소를 갖추고 있어 최상위요소와 핵심요소를 대상으로 메타데이터 요소를 설계하였다. 이 부분은 연구의 제한점으로 제시할 수 있다.

많은 분야에서는 해당 분야의 데이터를 관리 및 기술하기 위하여 각 분야에 특화된 메타데이터가 설계되어 활용되고 있지만, 임산공학에 대해선 거의 전무한 실정이다. 따라서 본 연구는 임산공학에서 생산된 데이터를 체계적으로 관리함으로써 국가와 기관의 자산인 연구데이

터를 보존 및 재사용할 수 있는 기반을 마련하고자 한다. 본 연구를 통해 제안되는 메타데이터 요소를 기반으로 연구데이터 관리 및 공유 시스템 구축 시 데이터베이스를 설계할 수 있으며, 외부 시스템과 연계시 메타데이터 교환 형식을 제시하는데 사용할 수 있을 것으로 판단된다.

2. 이론적 배경

본 절에서는 메타데이터 설계와 관련된 국내외 선행 연구를 분석하여 시사점을 도출하고자 한다. 다음의 〈표 1〉은 국내외 메타데이터 관련 연구를 요약한 것이다.

이미화, 이은주, 노지현(2020)은 국립중앙도서관의 오픈액세스 리포지토리인 OAK에서 연구데이터를 기술하기 위하여 기존 메타데이터의 확장 방안을 제안하였다. 해당 연구에서는 한국연구재단의 기초학문자료센터, 한국과학기술정보연구원의 국가연구데이터플랫폼, 샌디에고 대학 도서관 컬렉션 그리고 미시간 데이터 리포지토리를 분석하여 OAK의 연구데이터 기

1) ISO 19115는 16개의 패키지와 총 411개의 항목으로 정의된다.

〈표 1〉 국내 메타데이터 관련 선행연구 요약

제목	발행년도	비고
연구데이터 관리를 위한 OAK 메타데이터 확장 방안 연구	2020	국내
한의학 연구 데이터 관리 및 공유를 위한 메타데이터 요소 설계	2019	국내
ISO 19115-1 지리 메타데이터 표준과 CodeMeta와의 매핑	2019	해외
지구 관측 분야에서 효과적인 메타데이터 설계를 위한 방법론	2017	해외
HydroShare의 수문학 분야 애플리케이션을 통한 환경모델 메타데이터 프레임워크 설계	2017	해외
풍수해 피해예측지도 연계·활용을 위한 표준메타데이터 설계	2017	국내
침수지도 수집 및 관리를 위한 벡터형 공간정보 표준 기반의 메타데이터 설계	2016	국내
PREMIS 기반 보존 메타데이터 요소 개발에 관한 연구	2012	국내
메타데이터 애플리케이션 프로파일 개발을 위한 최신 방법론	2012	해외
활용도 제고를 위한 지적정보의 메타데이터 표준화 적용방안	2011	국내
산림 정보자원을 위한 메타데이터 스키마	2006	해외

술을 위한 메타데이터 확장 방안을 제시하였다. 연구 내용에는 기존 메타데이터에서 연구데이터 기술을 위한 새로운 메타데이터 요소를 추가하였으며 또한, 검색을 위하여 통제어휘집과 구문에 대한 사항도 제시하고 있다. 예상준, 장호, 김선태(2019)는 한의학 연구 분야 연구데이터 관리 및 공유를 위해 리포지터리에서 사용될 스키마를 제안하고자 최상위 요소 27개, 필수요소 29개, 권고요소 13개, 선택요소 31개를 메타데이터로 도출하였다. Habermann(2019)은 과학 연구용 소프트웨어의 인용 및 재사용을 목적으로 과학 소프트웨어와 코드를 위한 최소 메타데이터 스키마를 만드는 과학 및 산학 관계자들의 공동 프로젝트를 수행하였다(CodeMeta 2019). 이 코드메타 프로젝트에서는 소프트웨어 메타데이터를 위한 60개 이상의 어휘 매핑을 제안하였다. 제안된 매핑은 (1) 익숙한 통용어로 소프트웨어 메타데이터를 표현하고 (2) ISO 표준을 사용한 소프트웨어 메타데이터를 JSON-LD로 쉽게 변환할 수 있도록 하였다. Desconnets, Mougnot 그리고 Iahdi

(2017)는 지구 관측 데이터를 위해 설계된 어플리케이션 프로파일(EOAP)을 개발하는데 사용된 방법론적인 고려사항을 제안했다. EOAP 어플리케이션 프로파일은 환경 문제 분석을 위한 위성영상을 처리하기 위해 더블린 코어 어플리케이션 프로파일 가이드라인에 준수하여 설계되었다. 또한 일반 이용자를 대상으로 다중 센서, 다중 해상도 이미지 검색을 용이하도록 구조적이고 의미적인 상호운용 프레임워크를 제시하였다. 방법론은 주로 DCAM(Dublin Core Abstract Model), DSP(Description Set Profile) 구조 모델 그리고 싱가포르 프레임워크를 바탕으로 하였다. Morsy 등(2017)은 수문학 분야 웹 기반 협업 시스템인 HydroShare에서 환경 모델을 공유 및 재사용하기 위해 환경 모델 메타데이터 프레임워크를 설계하였다. 저자는 환경 모델 특성상 여러 다른 종류로 이루어진 데이터와 모델을 수용하기 위해 메타데이터 프레임워크는 유연하고 일반적이어야 한다고 지적하였다. 또한 상호운용성을 촉진하기 위해 더블린 코어 표준을 사용하여 일반적인

설명(dcterms)을 기술하고, 기본 더블린 코어 메타데이터를 자원별로 확장(hsterms)하였다.²⁾ 더블린 코어를 사용한 일반적인 기술을 하는 요소는 네임스페이스에 접두사로 dcterms를, 더블린 코어를 확장한 요소는 접두사로 hsterms를 사용하였다. 설계한 Model program과 Model Instance 자원의 메타데이터 요소를 RDF triples로 표현하였다. 서강현 등(2017)은 풍수해 피해예측지도 연계 및 활용의 고도화를 위해 ISO/TC211-19115를 참조 표준으로 홍우, 산사태, 바람에 대한 풍수해 피해예측지도 표준 메타데이터를 설계하고, 이를 기반으로 하는 표준 메타시스템을 시범적으로 구축해 결과를 검증하였다. 심규성(2016)은 국가재난정보의 통합적 운영 및 관리를 위해 표준화된 기준을 정립하고자 ISO/TC211-19115를 기반으로 유지보수, 제약, 메타데이터, 공간정보 체계, 인식, 배포별로 활용하기에 적절한 메타데이터 표준안을 도출하였다. 도출한 메타데이터의 UML과 코드 리스트를 작성하고 침수지도의 개체별 공간정보 데이터 샘플링 함으로써 속성정보 데이터베이스 설계와 메타데이터 연계 등 현업에서 활용할 수 있는 사례를 제시하였다. 박옥남(2012)은 PREMIS 프로젝트와 국립중앙도서관의 디지털 자원에 대한 분석을 바탕으로 보존 메타데이터 스키마를 개발하였다. 또한 MODS와 METS를 기반으로 한 국립중앙도서관의 기존 목록 레코드에 개발한 보존 메타데이터를 적용 및 활용할 수 있는 구체적인 가이드라인을 제시하였다. Malta와 Baptista(2012)는 메타데이터 어플리케이션 프로파일 개발을 위한 종합

적인 방법론 수립에 기여하기 위해 메타데이터 최신 방법론에 대한 정보를 수집하여 제시하였다. 정보 수집을 위해 어플리케이션 프로파일의 기존 개발 분석 및 식별, 기존 방법론 분석 및 식별이라는 2개의 부분으로 나누고 검색 키워드를 제시하였다. 일반 검색, 온라인 데이터베이스 검색, 메타데이터 어플리케이션 프로파일 검색을 수행하였고 검색 방법에 대해 상세히 서술하였다. 전방진(2011)은 지적정보의 메타데이터 표준화를 위하여 메타데이터 설계, 메타데이터 관리 및 확장성과 수준별 항목을 위한 업무별·사용자별 상세요구 분석, 최신 국제표준(ISO) 메타데이터반영과 지속적인 모니터링, 지적정보의 유통을 위한 관련 서비스 기술표준 분석 등의 사항들을 제시하였다. 이밖에 지적정보의 법적 보완 문제와 메타데이터 표준화의 법적 근거 마련, 체계적인 추진체계의 필요성을 제안하였다. 마지막으로 Schuck(2006)은 산림정보 시스템에서 사용될 산림 정보 자원의 메타데이터 설계를 위해 더블린 코어를 기반으로 메타데이터 스키마를 설계하였다. 단순 더블린 코어 15개 요소를 모두 수용하고 2개의 추가 요소(audience, reference system)를 제시하였다. 15개 요소 중 일부 요소의 하위요소를 전개하고, subject 요소의 하위 요소를 추가하는 과정에서 NEFIS 통제 어휘/키워드리스트가 확립되었다. 메타데이터 스키마 내에서 정보자원의 품질 보고를 구체화하기 위한 접근법으로 하위요소 quality report가 제안되었다.

이상과 같이 10건의 국내외 선행 연구를 살펴본 결과 국내 선행 연구의 경우 한의학, 지리

2) 더블린 코어를 사용한 일반적인 기술을 하는 요소는 네임스페이스에 접두사로 dcterms를, 더블린 코어를 확장한 요소는 접두사로 hsterms를 사용하였다.

정보, 도서관 등 특정 분야에 특화된 메타데이터 개발에 초점이 맞추어진 것을 확인할 수 있었다. 해외 선행 연구도 국내와 유사하게 지리 정보, 지구관측, 수문학 등 특정 분야에 적용할 수 있는 메타데이터 개발에 관한 논문을 확인할 수 있었다. 산림 분야와 유사한 논문으로 기술한바와 같이 2006년에 작성된 『산림 정보자원을 위한 메타데이터 스키마』만이 유일한 관련 논문으로 확인되었다. 관련 연구의 경우 최신성이 부족한 부분과 단순한 더블린 코어를 활용한 메타데이터 스키마라는 제한점이 있어 임산공학 분야에 적용하는 게 한계가 있다고 판단하였다. 따라서 본 논문에서는 기존에 개발된 지리 정보, 연구데이터 기술 등을 위한 표준 메타데이터를 활용하여 임산공학에 특화된 메타데이터를 개발하고자 한다.

3. 임산공학 분야 관련 메타데이터 요건 분석

본 장에서는 임산공학 분야를 위한 메타데이터를 개발하기 위하여 관련된 표준 메타데이터를 조사 및 분석하고자 한다. 관련된 표준 메타

데이터는 <표 2>와 같다.

임산공학 분야 메타데이터를 개발하기 위하여 조사 및 분석대상으로 제시한 메타데이터는 FGDC-CSDGM, ISO 19115, NEFIS 메타데이터 스키마, INSPIRE, ANZLIC 메타데이터 프로파일, TTAK.K0-10.0976 그리고 DataCite 메타데이터 스키마 4.3 등이다. 이 중 NEFIS(산림 분야)와 TTA, DataCite(연구데이터)를 제외하고 4개 표준은 지리 정보와 관련되어 있다. 지리 정보와 관련된 메타데이터를 선택한 이유는 임산 공학 특성상 지리 정보에 대한 메타데이터 항목을 기술해야 하므로 분석 대상에 포함하게 되었다. 또한 위에서 제시된 4개의 표준이 일반적으로 전 세계에서 지리정보에 가장 많이 활용되고 있는 표준이므로 본 연구에서 포함하게 되었다.

3.1 FGDC-CSDGM

CSDGM(Content Standards for Digital Geospatial Metadata)은 디지털 지형 공간 메타데이터를 위한 미국 연방 지리정보위원회 콘텐츠 표준이다. 1994년 4월, Executive Order 12906을 통한 국가 지리 공간 데이터 기반 구축

<표 2> 조사 및 분석 대상 메타데이터

메타데이터	분야	비고
FGDC-CSDGM	지리 정보	미국
ISO 19115	지리 정보	
NEFIS 메타데이터 스키마	산림 분야	유럽
INSPIRE	지리 정보	유럽
ANZLIC 메타데이터 프로파일	지리 정보	호주
TTAK,K0-10.0976	연구데이터	한국
DataCite 메타데이터 스키마 4.3	연구데이터	

(NSDI, National Spatial Data Infrastructure)에 따라 미국의 연방지리정보위원회(FGDC, US Federal Geographic Data Committee)에서 지리 공간 데이터의 공동이용과 배포 및 공유를 위해 개발하였다. CDSGM은 디지털 지리 공간 데이터의 메타데이터로서 지리 정보 메타데이터와 관련된 용어 및 개념에 대한 정의를 제공하고 있다. ISO 15046 Part 15와 호환성이 고려된 현재 버전(FDDC-STD-001-1998; CSDGM ver 2.0)은 1998년 7월에 발행되었고 널리 사용되었으나, 2010년 9월부터 FGDC에 의해 ISO 19115를 승인하고 연방 기관에서 ISO 표준으로 전환하도록 장려하였다. 또한 생물학, 해안선, 원격 감지를 지원하기 위하여 CSDGM 확장 프로파일도 개발되었다.³⁾

CSDGM 표준은 10개의 섹션(식별 정보, 데이터 품질정보, 공간데이터 조직 정보, 공간 참조 정보, 엔티티 및 속성 정보, 배포 정보, 메타데이터 참조 정보, 인용 정보, 기간 정보, 연락 정보)과 각 섹션 아래 하위요소 69개로 구성되어 있다.

식별 정보 섹션은 데이터셋에 대한 기본정보를, 데이터 품질정보 섹션은 데이터셋에 대한 일반적인 평가를 나타낸다. 공간데이터 조직 정보 섹션은 데이터셋 내의 공간정보를 표현하기 위해 사용된 메커니즘을 제시한다. 공간 참조정보 섹션은 데이터셋 내의 좌표를 위한 참조를 제공한다. 엔티티 및 속성 정보 섹션은 엔티티 유형, 속성 및 속성값이 할당될 수 있는 도메인 등을 포함한 데이터셋의 정보 내용에

대한 자세한 사항을 명시한다. 배포 정보 섹션은 데이터셋을 얻기 위한 옵션과 배포자에 대한 정보를 제공한다. 메타데이터 참조 정보 섹션은 책임 담당자 및 메타데이터 정보의 최신성에 대한 정보를 제공한다. 인용 정보 섹션은 데이터셋의 인용을, 기간 나타내는 수단을 제공한다. 기간 정보 섹션은 이벤트 날짜 및 시간에 대한 정보로 시간 정보를 나타내는 수단을 제공한다. 연락 정보 섹션은 데이터셋과 관련된 조직 및 개인을 식별하고 의사소통하기 위한 수단을 제공한다. 앞서 서술한 기간 인용 정보 섹션, 기간 정보 섹션, 연락 정보 섹션은 단독으로 사용되지 않고 메타데이터 표준의 다른 섹션에서 사용된다(The Federal Geographic Data Committee 2019).

3.2 ISO 19115

ISO(International Standardization Organization)는 1946년에 설립된 표준화를 위한 국제 기구이다. ISO 산하 지리 정보/지질학 기술 위원회인 ISO/TC211(ISO/TC 211 Geographic information/Geomatics)은 지리 정보에 대한 국제 표준을 담당한다. ISO/TC211은 2003년에 지리 정보 및 서비스를 기술하기 위해 ISO 19115:2003 Geographic information - Metadata를 발표하였다. ISO 19115는 디지털 지리 데이터 및 서비스의 식별, 범위, 품질, 공간 및 시간적 측면, 내용, 공간적 참조, 묘사, 배포 및 기타 특성에 대한 정보를 제공한다(ISO 2014).

3) CSDGM 확장 프로파일로는 CSDGM: BIOLOGICAL DATA PROFILE(생물학적 데이터 프로파일), CSDGM: METADATA PROFILE FOR SHORELINE DATA(해안선 데이터를 위한 메타데이터 프로파일), CSDGM: Extensions for Remote Sensing Metadata(원격 감지 메타데이터 확장) 등이 있다.

현재, 지리정보, 해양, 기상 분야 등 여러 분야에서 ISO 19115를 이용한 확장 메타데이터를 사용하고 있다. 지리정보 분야에는 호주·뉴질랜드의 ANZLIC Metadata Profile, 북미의 ISO 19115 North American Profile, 영국의 UK AGMAP, Gemini, 일본의 JPGIS 등이 해당한다. 해양 분야에는 The National Oceanographic Data Centre의 Cruise Summary Reports, 유럽의 EDMED Metadata Profile, 호주의 Marine Community Profile 등이 존재한다. 그 외에도 기상 분야의 WMO Core Metadata Profile 등이 있다(DCC 2019).

ISO 19115는 16개의 패키지와 총 411개의 항목으로 정의되며 ISO 19115의 핵심 요소는 22(필수: 7 / 조건부: 4 / 선택: 11)개이다. 필수 요소에는 메타데이터 연락처정보, 메타데이터 생성 일자, 데이터셋 제목, 데이터셋 참조 일자, 데이터셋 요약설명, 데이터셋 언어, 데이터셋 주제 범주 등이 해당한다. 선택요소로는 메타데이터 언어, 메타데이터 문자셋, 데이터셋 문자셋 그리고 데이터셋의 지리적 위치 등 7개 항목이 포함된다. 조건부 요소에는 메타데이터 언어, 메타데이터 문자셋, 데이터셋 문자셋 그리고 데이터셋의 지리적 위치 등 4개 항목이 있다. 마지막으로 선택 요소에는 메타데이터 파일 식별자, 메타데이터 표준명, 메타데이터 표준 버전, 데이터셋 책임담당자, 공간 표현 유형, 데이터셋의 공간해상도, 데이터셋의 부가적인 범위정보, 이력, 참조 체계, 배포 포맷 그리고 온라인 자원 등 11개의 요소가 포함되어 있다(NOAA Environmental Data Management Wiki 2010).

3.3 NEFIS 메타데이터 스키마

NEFIS(Network for a European Forest Information Service) 프로젝트는 유럽 내 다양한 유형의 산림정보의 처리 및 배포를 향상하기 위해 임업 및 정보기술 개발 분야의 중요 주체와 이해관계자를 포함한 네트워크를 구축하는 프로젝트이다. 2003년부터 2005년까지 유럽공동체 위원회의 Quality of Life and Management of Living Resources Programme의 동반조치로 European Forest Institute에 의해 수행되었다. 이 프로젝트는 산림 정보 시스템을 구축하여 지속적인 데이터와 DB의 가치를 극대화하고 서로 다른 인터넷 기반 산림 정보 서비스간의 사소통을 위하여 통일된 운영 방법을 개발하기 위하여 시작되었다(NEFIS 2019).

NEFIS 메타데이터 스키마는 정보자원의 식별 및 목록을 위한 메타데이터에 중점을 두고 더블린코어, ISO 19115 표준을 기반으로 개발하였다. 더블린코어의 15개 필수요소와 1개의 선택요소, ISO 19115의 1개 요소, 자체 추가 요소 1개로 되어있다. 이 중 2개의 더블린 코어 필수요소 아래 추가로 하위요소를 전개하였다. 즉, 더블린코어의 15개 요소가 모두 사용되었으며, 더블린코어 선택요소인 '이용대상(Audience)'와 ISO19115를 참조한 '참조체계(Reference system)' 그리고 '접근 권한(Access right)'을 새로 추가하였다. 기존 더블린코어 요소인 'Subject' 아래 상세구분으로 'NEFIS테마(NEFISthemes)', 'NEFIS용어(NEFISterms)', '후보용어(Nominatedterms)'를, '설명(Description)'요소 아래에 '품질보고(Quality report)'를 추가하였다(Schuck and Andreas 2007).

3.4 INSPIRE

INSPIRE(Infrastructure for Spatial Information in the European Community) 지침(directive)⁴⁾은 유럽 전역의 환경 정책을 목적으로 하는 공간 데이터 인프라 구축을 위해 2007년 5월 15일 발효된 법령⁵⁾이다(wikipedia 2019). INSPIRE 지침은 6가지 의무적인 실행 규칙(Implementing Rules)을 가지며 이 중 메타데이터가 포함되어 있다. 지침에 따라 EU 회원국은 공간 데이터셋 및 서비스에 대한 메타데이터를 작성하고 이 메타데이터를 최신 상태로 유지하도록 보장해야 한다. INSPIRE의 메타데이터에 대한 실행규칙과 기술 가이드라인은 메타데이터의 생성 및 유지보수에 대한 요건을 규정한다. 해당 실행규칙⁶⁾은 INSPIRE 지침을 준수하는데 필요한 최소 메타데이터 요소 집합을 제시하며 법적 구속력을 가진다. 기술 가이드라인(Technical Guidelines)은 법적 구속력은 없으나 법적 의무를 구현할 방법을 명시한다. 기술가이드라인으로는 ISO 19115 및 ISO 19119 기반 기술 가이드라인, ISO/TS 19139 기반 INSPIRE 데이터셋 및 서비스 메타데이터 기술 가이드라인을 제공한다.

INSPIRE 실행 규칙에서 제시하는 메타데이터 요소는 INSPIRE 지침을 준수하는데 필요한 최소 메타데이터 요소 집합으로, 회원국의 필요에 따라 확장해서 사용할 수 있다. INSPIRE 실행 규칙에서 제시하는 최소 메타데이터 요소는 식별, 공간 데이터 및 서비스의 분류, 지리적 위치, 시간 참조, 품질과 유효성, 적합성, 접근 및 이용에 대한 제한조건, 구축·관리·유지 담당 기관, 메타데이터에 대한 메타데이터 등 10개의 범주로 나눌 수 있으며 총 27개 요소가 포함되어 있다(European Commission 2008).

3.5 ANZLIC 메타데이터 프로파일

ANZLIC(The Australia New Zealand Land Information Council)은 공간 정보를 담당하는 호주와 뉴질랜드의 최고 정부 기관이다. ANZLIC은 공간 정보의 접근성과 사용성을 촉진하기 위한 정책과 전략을 개발하는 역할을 한다. ANZLIC은 2001년 오스트레일리아와 뉴질랜드를 위한 메타데이터 가이드라인⁷⁾을 개발하였다. 그 이후 ISO/TC211을 통해 지리 정보 메타데이터 국제표준인 ISO19115:2003이 발표되었고 이 표준을 검토하여 호주와 뉴질랜드 국가표준으

4) directive: 지침(directive)은 유럽연합의 법률조치로서, 결과를 달성하기 위한 수단을 지시하지 않고 유럽연합 회원국들이 특정한 결과를 얻도록 요구하는 것이다. 스스로 실행하는 자율성을 갖는 규정과 달리 지침은 채택해야 할 정확한 규칙을 제시한다. 지침 초안은 유럽위원회 및 국가 전문가와 협의하여 작성되며 평가를 위해 회원국의 장관들로 구성된 의회와 위원회에 제출된다. - Directive (European Union) [Website]. Retrieved from [https://en.wikipedia.org/wiki/Directive_\(European_Union\)](https://en.wikipedia.org/wiki/Directive_(European_Union)).

5) Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).

6) COMMISSION REGULATION (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata.

7) ANZLIC Metadata Guidelines: Core metadata elements for geographic data in Australia and New Zealand (version 2, February 2001).

로 AS/NZS ISO19115:2005를 채택하였다. 이를 기반으로 호주와 뉴질랜드를 위한 메타데이터 프로파일을 개발하였다.

ANZLIC 메타데이터 프로파일은 AS/NZS ISO 19115:2005의 모든 요소를 채택하였다. 예외적으로, 메타데이터 요소 중 파일 식별자(file Identifier)는 필수요소로 지정하였다. 파일 식별자를 사용하여 메타데이터 레코드를 고유하게 식별할 수 있으며, parentIdentifier와 fileIdentifier를 사용하여 메타데이터 레코드 간의 하위-부모 관계를 식별할 수 있도록 하였다. 또한, ANZLIC 메타데이터 프로파일은 지리 데이터셋에 권장되는 핵심 메타데이터를 정의하고 있는데 핵심 메타데이터에는 메타데이터 요소의 최소 집합과 탐색을 위한 지리 데이터셋 기술(description)을 향상시키는 ANZLIC에 의해 식별된 추가 요소를 포함하고 있다.

3.6 TTA.K0-10.0976

범용적 연구 데이터 기술을 위해 수집된 국내 표준 메타데이터는 정보통신단체표준인 'TTA.K0-10.0976'이다. TTA.K0-10.0976은 연구 과제 수행 도중 생산된 연구 데이터를 메타데이터 기반으로 효율적으로 관리, 공유 및 활용 할 수 있도록 메타데이터 관리체계, 요소 및 세부 사항을 표준으로 제정한 것으로 해당 표준은 메타데이터 프로젝트그룹(PG606)에 제안되고, 소프트웨어/콘텐츠 기술위원회(TC6)에서 심의를 통과하여 2017년 3월에 승인된 정보통신단체표준이다(한국정보통신기술협회 2017). 또한 TTA.K0-10.0976은 국내의 정보통신단체표준으로 승인된 것으로 연구

데이터를 효과적으로 관리, 공유하고 재활용함으로써 연구의 효율성을 제고하는 도구로 활용될 목적으로 제안되었다(김선태 2019).

해당 표준은 '컬렉션 메타데이터', '데이터셋 메타데이터', '파일 메타데이터' 그리고 '리포지터리 메타데이터'로 구성되어 있어 있는데 데이터를 기술하기 위해 필요한 데이터셋 메타데이터는 15개 항목을 포함하고 있다. 적용 수준으로 데이터셋 메타데이터 표준을 분석하면 '사용 필수'에 포함되는 메타데이터 요소는 'Identifier', 'Title', 'Creator', 'Publisher' 그리고 'Publicationyear' 등 5개이다. 다음 '사용 권고'에 해당하는 메타데이터 요소는 'Date', 'Description', 'Subject', 'Contributor', 'Contact' 그리고 'Rights' 등 6개 항목이다. 마지막으로 '사용 선택' 항목에는 'Dataset', 'Keyword', 'AccessType' 그리고 'AccessRestriction' 등 4개 요소가 포함되었다.

3.7 DataCite 메타데이터 스키마 4.3

마지막으로 조사 및 분석한 메타데이터는 국외 사실적 표준으로 범용적 연구 데이터 기술을 위한 표준 메타데이터인 'DataCite'이다. DataCite는 연구 데이터 및 기타 연구 결과물에 DOI(Digital Object Identifier)와 같은 영구 식별자를 제공하는 세계적인 비영리 단체이며, 연구 커뮤니티 내의 조직은 회원으로 DataCite에 가입하여 모든 연구 결과물에 DOI를 할당 할 수 있다. 이러한 DataCite는 데이터셋의 접근, 식별, 공유 및 재사용에 대한 지속적으로 접근하기 위하여 2009년 말에 설립되었다(DataCite 2019). DataCite 메타데이터 스키마는 인용 및 검색을

위해 정보자원을 정확하고 일관되게 식별하도록 선택된 핵심 메타데이터 Property 목록이다.

DataCite 메타데이터 스키마 4.3의 최상위 요소는 필수 요소 6개, 권고 요소 6개 그리고 선택 요소 7개로 구성되어 있다. 필수 요소에는 'Identifier', 'Creator', 'Title', 'Publisher', 'PublicationYear' 그리고 'ResourceType' 등 6개의 항목을 포함하고 있다. 다음으로 권고 요소는 'Subject', 'Contributor', 'Date', 'RelatedIdentifier', 'Description' 그리고 'GeoLocation' 등 6개의 항목을 가지고 있다. 마지막으로 선택 요소에는 'Language', 'AlternateIdentifier', 'Size', 'Format', 'Version', 'Rights' 그리고 'FundingReference' 등 7개가 있다.

4. 임산공학 분야 연구데이터 관리를 위한 메타데이터 요소 선정

이번 장에서는 3장에서 분석한 7개의 메타데이터 표준을 기준으로 임산공학을 위한 메타데

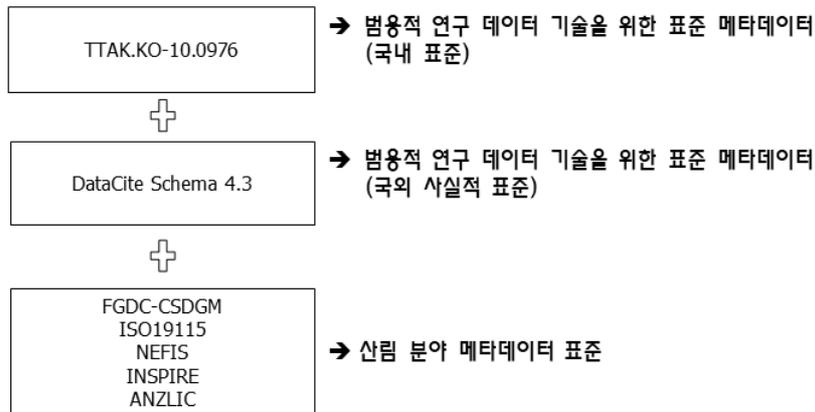
이터 항목을 도출하였다.

4.1 임산공학 분야를 위한 메타데이터 요소 도출을 위한 선정 기준

임산공학 분야의 메타데이터를 관리하기 위한 요소를 도출하기 위하여 본 연구에서는 3장에서 분석한 7개의 메타데이터를 크로스워크(Crosswalk)하였다. 여기에서 말하는 크로스워크란, 여러 메타데이터 간에 메타데이터 요소의 의미와 구조를 매핑하는 것으로 메타데이터간의 상호운용성을 확보하기 위하여 자주 사용되는 방법이다(고영만, 서태설, 임태훈 2007).

다음의 <그림 2>는 이번 연구의 핵심인 임산공학 분야의 메타데이터를 도출하기 위한 사용된 메타데이터 항목을 도식화한 것이다.

먼저, 임산공학 분야의 메타데이터 요소를 도출하기 위하여 TTA와 DataCite의 요소를 분석 기준으로 활용하였다. 다음의 <표 3>은 분석 기준으로 도출한 메타데이터 요소이다.



<그림 2> 임산공학 분야 메타데이터 개발 과정

〈표 3〉 임산공학 분야 메타데이터 도출을 위한 기준 요소

메타데이터 요소	메타데이터 의미	적용 수준	비고
Identifier (식별자)	데이터셋의 식별자 또는 식별값	M	TTA DataCite
Title (제목)	데이터셋의 이름 또는 명칭	M	TTA DataCite
Creator (생산자)	데이터셋을 생성 또는 묶은 사람, 조직 등을 기술함	M	TTA DataCite
Publisher (출판자)	데이터셋을 웹에 공개 또는 출판하는 주체	M	TTA DataCite
Publication Year (출판년도)	데이터셋이 웹상에서 접근 또는 이용 가능해진 연도	M	TTA DataCite
Date (날짜)	데이터셋이 생애주기 동안 발생한 다양한 이벤트 날짜	R	TTA DataCite
Language (언어)	정보자원의 기본 언어	O	DataCite
Description (데이터셋 설명)	데이터셋의 내용에 대한 설명	R	TTA DataCite
Subject (주제)	데이터셋의 내용이 지닌 주제 정보를 기술함	R	TTA DataCite
Contributor (기여자)	데이터셋의 수집, 생산, 예산지원 등에 관계된 사람이나 기관 등을 기술함	R	TTA DataCite
Rights (저작권)	데이터셋의 저작권 정보를 기술함	O	TTA DataCite
Access Restriction (접근 제한)	데이터셋의 접근제한 또는 이용제한 유형	O	TTA
Resource Type (데이터셋 유형)	데이터셋 유형을 자유롭게 기술	M	DataCite
AlternateIdentifier (대체식별자)	등록 중인 자원에 적용된 1차 식별자 이외의 식별자	O	DataCite
RelatedIdentifier (관련식별자)	관련 자원의 식별자	R	DataCite
Funding Reference (펀딩 참조정보)	연구비 정보로서 연구비 지원기관 정보, 연구비, 프로젝트명, 프로젝트 코드 등을 기술	O	DataCite
Format (파일 형식)	파일 형식을 확장자명과 간단한 설명으로 기입함	O	DataCite
Version (버전)	정보자원의 버전 번호	O	DataCite
Size (파일 크기)	파일의 크기(용량)를 단위 없이 값만 기술함	O	DataCite
Geographic Location (위치 정보)	데이터가 수집되었거나 데이터가 집중된 공간 영역 또는 명명된 장소	R	DataCite

〈표 3〉은 TTA, K0-10.0976와 DataCite 메타데이터 스키마 4.3을 기준으로 도출한 20개의 메타데이터 요소와 의미이다. TTA 표준의 경우, 15개 항목 중 메타데이터를 도출하기 위한 기준에 적합하지 않은 'Dataset', 'Contact', 'Keyword' 그리고 'AccessType'은 제외하였다. 제외한 이유는 해당 요소들이 컬렉션을 기술하기 위한 메타데이터 요소이기 때문이다. 여기에 DataCite에서 가져온 'Language', 'ResourceType', 'AlternateIdentifier', 'RelatedIdentifier', 'Funding Reference', 'Format', 'Version', 'Size' 그리고 'Geographic Location' 등 9개의 요소를 포함하였다.

먼저, TTA와 DataCite 둘 다에 포함된 요소는 'Identifier', 'Title', 'Creator', 'Publisher', 'Publication Year', 'Date', 'Description', 'Subject', 'Contributor' 그리고 'Rights' 등 10개 항목이다. TTA에만 포함된 요소는 'Access Restriction'이며, DataCite에만 포함된 요소는 'Language', 'Resource Type', 'AlternateIdentifier', 'RelatedIdentifier', 'Funding Reference', 'Format', 'Version', 'Size' 그리고 'Geographic Location' 등 10개이다.

다음 적용수준으로 분석한 결과, 'M(필수)'에 속하는 요소는 'Identifier', 'Title', 'Creator', 'Publisher', 'Publication Year' 그리고 'Resource Type' 등 6개이며 'R(권고)'에 속하는 요소는 'Date', 'Description', 'Subject', 'Contributor', 'RelatedIdentifier' 그리고 'Geographic Location' 등 6개이다. 마지막으로 'O(선택)'에는 'Language', 'Rights', 'Access Restriction', 'AlternateIdentifier', 'Funding Reference', 'Format', 'Version' 그리고 'Size' 등 8개를 포함하고 있다.

4.2 임산공학 분야를 위한 메타데이터 요소 도출

다음의 〈표 4〉는 3장에서 분석한 임산공학 분야의 메타데이터를 〈표 3〉에서 도출한 기준으로 크로스워크를 진행한 것이다.

크로스워크를 진행하기 위한 메타데이터에는 ISO 19115, NEFIS, INSPIRE 그리고 ANZLIC 등 4개가 포함되었다. 3장에서 분석한 FGDC-CSDGM이 크로스워크 대상에서 제외된 이유는 FGDC-CSDGM이 추후 ISO 19115로 전환되었기 때문이다.

제시된 메타데이터 중 'ISO19115'는 핵심 요소 집합을 'NEFIS'는 전체 메타데이터 스키마를 'INSPIRE'는 실행규칙 메타데이터 요소를 마지막으로 'ANZLIC'은 핵심 메타데이터 집합을 크로스워크 대상으로 삼았다.

분석결과 모든 메타데이터에 포함된 요소는 'Title', 'Creator', 'Date', 'Language', 'Description', 'Subject' 그리고 'Geographic Location' 등 7개이다. TTA/DataCite 기준에는 포함되어 있지 않지만, 4개의 메타데이터에 모두 포함된 요소는 'Lineage(NEFIS의 경우, source를 지칭)'이다. 다음 TTA/DataCite 기준으로 3개의 메타데이터에 속한 요소는 'Identifier' 1개 항목이며, 2개의 메타데이터에 포함된 요소는 'Resource Type', 'Metadata of Metadata' 그리고 'Access Right' 등 3개이다. 여기에서 'Metadata of Metadata'는 메타데이터를 기술하기 위한 메타데이터와 관련된 항목이면 해당 항목에 모두 포함시켰다. 또한, TTA/DataCite 기준으로 1개의 메타데이터에 출현한 항목은 'Publisher', 'Publication Year', 'Contributor',

〈표 4〉 임산공학 분야 메타데이터 요소 도출을 위한 크로스워크(Crosswalk)

TTA/DataCite	ISO19115	NEFIS	INSPIRE	ANZLIC
Identifier		Identifier	<ul style="list-style-type: none"> • Unique resource identifier • Resource locator • Coupled resource 	Dataset title
Title	Dataset title	Title	Resource title	Dataset title
Creator	Dataset responsible party	Creator	<ul style="list-style-type: none"> • Responsible party • Responsible party role 	Dataset responsible party
Publisher		Publisher		
Publication Year			Date of publication	
Date	Dataset reference date	<ul style="list-style-type: none"> • Date • Coverage 	<ul style="list-style-type: none"> • Date of creation • Temporal extent • Date of last revision 	<ul style="list-style-type: none"> • Dataset reference date • Temporal extent information for the dataset
Language	<ul style="list-style-type: none"> • Dataset language • Metadata character set 	Language	Resource language	Dataset language
Description	Abstract describing the dataset	<ul style="list-style-type: none"> • Description • Relation 	<ul style="list-style-type: none"> • Resource abstract • Keyword value • Originating controlled vocabulary 	<ul style="list-style-type: none"> • Abstract describing the dataset
Subject	Dataset topic category	Subject	<ul style="list-style-type: none"> • Topic category • Spatial data service type 	<ul style="list-style-type: none"> • Dataset topic category • Spatial representation type
Contributor		Contributor		
Rights		Rights		
Access Restriction			Limitations on public access	
Resource Type		Type	Resource type	
AlternateIdentifier				
RelatedIdentifier				
Funding Reference				
Format		Format		
Version				
Size				
Geographic Location	<ul style="list-style-type: none"> • Geographic location of the dataset • Additional extent information for the dataset • Spatial resolution of the dataset • Reference system • Spatial representation type 	<ul style="list-style-type: none"> • Reference system • Coverage 	<ul style="list-style-type: none"> • Geographic bounding box • Spatial resolution • Geographic bounding box • Spatial resolution 	<ul style="list-style-type: none"> • Spatial resolution of the dataset • Additional extent information for the dataset • Reference system • Vertical extent information for the dataset
	Lineage	Source	Lineage	Lineage
	Distribution format			Distribution format
	On-line resource			On-line resource
	<ul style="list-style-type: none"> • Metadata file identifier • Metadata language • Metadata character set • Metadata point of contact • Metadata date stamp • Metadata standard name • Metadata standard version 		<ul style="list-style-type: none"> • Metadata point of contact • Metadata date • Metadata language 	
		Audience		
		Access Right	Conditions applying to access and use	
			Specification	
			Degree	

'Rights', 'Access Restriction' 그리고 'Format' 등 6개이다. 마지막으로 TTA/DataCite에만 포함되었거나, TTA/ DataCite에 포함되지 않았지만 타 메타데이터 중 1개의 메타데이터에만 속한 항목은 'AlternateIdentifier', 'RelatedIdentifier', 'Funding Reference', 'Version', 'Size', 'Audience',

'Specification' 그리고 'Degree' 등 8개 요소이다.

TTA/DataCite를 기준으로 임산공학 분야의 4개 메타데이터를 분석한 결과, 다음의 <표 5>와 같이 임산공학을 위한 메타데이터 요소가 도출되었다.

<표 5> 임산공학을 위한 메타데이터 요소 도출

No.	메타데이터 요소	메타데이터 의미
1	Identifier(식별자)	데이터셋의 식별자 또는 식별값
2	Title(제목)	데이터셋의 이름 또는 명칭
3	Creator(생산자)	데이터셋을 생성 또는 묶은 사람, 조직 등을 기술함
4	Publisher(출판자)	데이터셋을 웹에 공개 또는 출판하는 주체
5	Publication Year(출판년도)	데이터셋이 웹상에서 접근 또는 이용 가능해진 연도
6	Date(날짜)	데이터셋이 생애주기 동안 발생한 다양한 이벤트 날짜
7	Language(언어)	정보자원의 기본 언어
8	Description(데이터셋 설명)	데이터셋의 내용에 대한 설명
9	Subject(주제)	데이터셋의 내용이 지닌 주제 정보를 기술함
10	Contributor(기여자)	데이터셋의 수집, 생산, 예산지원 등에 관계된 사람이나 기관 등을 기술함
11	Rights(저작권)	데이터셋의 저작권 정보를 기술함
12	Access Restriction(접근 제한)	데이터셋의 접근제한 또는 이용제한 유형
13	Resource Type(데이터셋 유형)	데이터셋 유형을 자유롭게 기술
14	AlternateIdentifier(대체식별자)	등록 중인 자원에 적용된 1차 식별자 이외의 식별자
15	RelatedIdentifier(관련식별자)	관련 자원의 식별자
16	Funding Reference (펀딩 참조정보)	연구비 정보로서 연구비 지원기관 정보, 연구비, 프로젝트명, 프로젝트 코드 등을 기술
17	Format(파일 형식)	파일 형식을 확장자명과 간단한 설명으로 기입함
18	Version(버전)	정보자원의 버전 번호
19	Size(파일 크기)	파일의 크기(용량)를 단위 없이 값만 기술함
20	Geographic Location(위치 정보)	데이터가 수집되었거나 데이터가 집중된 공간 영역 또는 명명된 장소
21	Lineage(이력)	데이터셋을 생성하는데 사용된 소스 및 처리 단계에 대한 기록
22	Distribution format(배포 형식)	배포되는 자원(공간정보)의 형식정보
23	On-line resource (온라인 정보자원)	데이터셋, 명세, 혹은 단체 프로파일명과 확장된 메타데이터 요소가 획득될 수 있는 온라인 자원에 대한 정보
24	Metadata of Metadata (메타데이터의 메타데이터)	메타데이터를 설명한 데이터
25	Audience(이용 대상)	해당 정보자원을 유용하게 사용할 수 있는 이용 대상
26	Access Right(접근 권한)	자원의 접근(이용) 권한이나 보안 상태에 대한 정보
27	Specification(규격)	데이터 형식 또는 연구 커뮤니티 요구사항 등과 관련된 규격
28	Degree(적합성 정도)	데이터 형식 및 내용에 요구되는 규격 준수 여부

다음의 <표 5>와 같이 TTA/DataCite와 임산공학 분야 메타데이터를 크로스워크 작업을 진행하여 도출된 메타데이터는 28개 항목이다. <표 4>에서 도출된 20개 항목에 'Lineage', 'Distribution format', 'On-line resource', 'Metadata of Metadata', 'Audience', 'Access Right', 'Specification' 그리고 'Degree' 등 8개 항목이 추가되었다. 즉, 8개 항목은 임산공학 분야와 관련된 고유의 메타데이터 요소라 할 수 있다.

4.3 임산공학 분야를 위한 메타데이터 요소 검증

이번 절에서는 4장 2절을 통해 도출된 메타데이터 요소를 검증하기 위해 국립산림과학원 소속 연구자를 대상으로 설문조사를 실시하였다. 설문조사 실시하기 전 임산공학 분야의 전문가인 국립산림과학원 연구자를 대상으로 메타데이터에 대한 교육을 실시하였다. 또한, 설문조사 결과를 반영하여 메타데이터 요소의 우선순위를 결정하였다.

설문 조사의 목적은 기 도출된 임산공학 분야 연구데이터를 기술하기 위한 메타데이터 항목의 중요도를 파악하기 위함이다. 설문 조사는 국립산림과학원 소속 연구자 19명을 대상으로 하였으며, 항목의 중요도를 파악하기 위해 5점 척도(매우 불필요, 불필요, 선택, 필수, 매우 필수)를 사용하였다.

다음은 설문 조사의 개요를 나타낸 것이다.

- 조사 목적: 임산공학 분야 메타데이터 항목 중요도 검증

- 조사 대상: 국립산림과학원 소속 연구자 19명
- 조사 일정: 2019년 09월 20일 ~ 2019년 10월 07일(18일간)
- 조사 방법: 웹 설문조사
(<http://ksurv.kr/?a=26461>)

다음 <표 6>은 설문응답자의 설문 결과를 바탕으로 각 항목에 대해 매우 불필요요소 (1점), 불필요요소 (2점), 선택요소 (3점), 필수요소 (4점), 매우 필수요소 (5점) 등으로 점수를 부여하여 순위화한 것이다.

다음의 <표 6>과 같이 가장 높은 점수를 얻은 메타데이터 요소는 'Title'으로 나타났다. 다음으로 2위부터 5위까지는 'Subject', 'Publication Year', 'Date', 'Resource Type'으로 나타났다. 6위부터 10위에 'Creator', 'Identifier', 'Format', 'Geographic Location', 'Description'이 해당한다. 11위부터 15위까지는 'Related Identifier', 'Version', 'Publisher', 'Rights', 'Lineage'로 나타났다. 다음으로 'Alternate Identifier', 'Contributor', 'Access Right', 'Metadata of metadata', 'Language'가 16위부터 20위를 차지하였다. 마지막으로 21위부터 28위는 'Size', 'On-line Resource', 'Access Restriction', 'Specification', 'Degree', 'Distribution Format', 'Funding Reference', 'Audience' 순으로 나타났다.

4.4 임산공학 분야를 위한 메타데이터 요소 재선정

이번 절에서는 4장 3절에서 도출된 메타데이터 항목에 우선순위 선정 기준을 부여하여 메

〈표 6〉 설문응답자의 메타데이터 요소 점수 부여 결과(5점 척도)

세부질문	매우 불필요 요소	불필요 요소	선택 요소	필수 요소	매우 필수 요소	점수 합계
데이터 이름(Title)	0	0	0	8	11	87
데이터 주제(Subject)	0	0	1	8	10	85
데이터 출판년도(Publication Year)	0	0	1	9	9	84
데이터 생성날짜(Date)	0	1	2	6	10	82
데이터셋 유형(Resource Type)	0	0	3	10	6	79
데이터 생산자(Creator)	0	0	6	7	6	76
식별자(Identifier)	1	0	5	6	7	75
데이터 형식(Format)	0	0	5	10	4	75
데이터 위치 정보(Geographic Location)	0	2	5	7	5	72
데이터 설명(Description)	0	1	6	8	4	72
관련 자원 식별자(Related Identifier)	1	1	5	7	5	71
데이터 버전(Version)	0	1	8	5	5	71
데이터 출판자(Publisher)	0	1	6	9	3	71
데이터 저작권(Rights)	0	2	4	10	3	71
데이터 이력 정보(Lineage)	0	2	5	10	2	69
대체 식별자(Alternate Identifier)	0	1	10	4	4	68
데이터 기여자(Contributor)	0	1	9	6	3	68
데이터 접근 권한(Access Right)	1	1	6	8	3	68
메타데이터를 설명한 데이터 (Metadata of Metadata)	0	1	8	8	2	68
데이터 기술 언어(Language)	0	1	10	5	3	67
데이터 크기(Size)	0	1	11	4	3	66
온라인 자원(On-line resource)	0	2	10	5	2	64
접근 제한(Access Restriction)	1	3	8	4	3	62
규격(Specification)	1	2	9	6	1	61
적합성 정도(Degree)	1	1	11	5	1	61
데이터 제공 형식(Distribution format)	0	4	10	4	1	59
펀딩 참조정보(Funding Reference)	0	3	12	4	0	58
이용 대상(Audience)	1	4	10	4	0	55

타데이터 요소를 필수, 권고 그리고 선택항목으로 재선정하였다. 우선순위 선정 기준은 다음과 같다.

- 기준 1: TTA, KO-10.0976 표준과 DataCite

Schema 4.3에서 제시한 필수 요소와 권고 요소는 본 연구에서 동일하게 필수 요소와 권고 요소로 한다.⁸⁾

- 기준 2: 기준 1의 권고요소에 해당하는 항목이 〈표 6〉의 최소 점수와 동일하거나 높

8) DataCite는 국제적으로 적용되는 사실적 표준으로 이번 연구에서는 메타데이터의 교환 및 상호운용성을 위하여 해당 표준에서 제시하는 필수 및 권고 항목은 그대로 적용하고자 한다.

은 경우 해당 항목은 모두 권고 요소로 한다(단, 이미 기준 1에서 제시한 요소의 적용수준은 변하지 않는다).⁹⁾

• 기준 3: 그 외의 요소는 선택 요소로 규정함

위에서 제시한 우선순위 선정 기준에 따라 다음의 <표 7>과 같이 28개 항목에 대한 우선 순위가 결정되었다. 28개 항목에는 필수 요소 6개, 권고 요소 13개 그리고 선택 요소 9개가 포함되었다.

'M(필수)'에 속하는 요소는 'Title', 'Publication Year', 'Resource Type', 'Creator', 'Identifier' 그리고 'Publisher' 등 6개이며 'R(권고)'에 속하는 요소는 'Subject', 'Date', 'Format', 'Description', 'Geographic Location', 'Rights', 'Version', 'Related Identifier', 'Lineage', 'Alternate Identifier', 'Metadata of Metadata', 'Access Right' 그리고 'Contributor' 등 13개이다. 'O

(선택)'에 속하는 요소는 'Language', 'Access Restriction', 'Size', 'On-line resource', 'Specification', 'Degree', 'Distribution format', 'Funding Reference' 그리고 'Audience' 등 9개이다.

위의 28개 요소 중 연구데이터 관리, 출판 및 인용하기 위한 메타데이터 요소는 'Title', 'Subject', 'Publication Year', 'Date', 'Resource Type', 'Creator', 'Identifier', 'Format', 'Description', 'Geographic Location', 'Rights', 'Version', 'Related Identifier', 'Publisher', 'Alternate Identifier', 'Contributor', 'Language', 'Access Restriction', 'Size' 그리고 'Funding Reference' 등 20개이다. 여기에 추가된 8개 요소는 임산공학 분야와 관련된 고유의 메타데이터 요소로서 'Lineage', 'Metadata of Metadata', 'Access Right', 'On-line resource', 'Specification', 'Degree', 'Distribution format' 그리고 'Audience'가 포함되었다.

<표 7> 임산공학 분야를 위한 메타데이터 요소

메타데이터 요소	메타데이터 의미	적용 수준	비고
Title (데이터 이름)	데이터셋의 이름 또는 명칭	M	
Subject (데이터 주제)	데이터셋의 내용이 지닌 주제 정보를 기술함	R	
Publication Year (데이터 출판년도)	데이터셋이 웹상에서 접근 또는 이용 가능해진 연도	M	
Date (데이터 생성 날짜)	데이터셋이 생애주기 동안 발생한 다양한 이벤트 날짜	R	
Resource Type (데이터셋 유형)	데이터셋 유형을 자유롭게 기술	M	
Creator (데이터 생산자)	데이터셋을 생성 또는 묶은 사람, 조직 등을 기술함	M	

9) 즉, 데이터 기여자(Contributor)는 기준 1에서 권고요소이다. 이 요소의 점수는 <표 6>에서 68점이며, 동일한 점수에 해당하는 요소는 데이터 이력 정보(Lineage), 대체 식별자(Alternate Identifier), 데이터 기여자(Contributor), 데이터 접근 권한(Access Right) 그리고 메타데이터를 설명한 데이터(Metadata of Metadata) 등 5개로 동일하게 권고요소로 간주한다.

메타데이터 요소	메타데이터 의미	적용 수준	비고
Identifier (식별자)	데이터셋의 식별자 또는 식별값	M	
Format (데이터 형식)	파일 형식을 확장자명과 간단한 설명으로 기입함	R	
Description (데이터 설명)	데이터셋의 내용에 대한 설명	R	
Geographic Location (데이터 위치 정보)	데이터가 수집되었거나 데이터가 집중된 공간 영역 또는 명명된 장소	R	
Rights (데이터 저작권)	데이터셋의 저작권 정보를 기술함	R	
Version (데이터 버전)	정보자원의 버전 번호	R	
Related Identifier (관련 자원 식별자)	관련 자원의 식별자	R	
Publisher (데이터 출판자)	데이터셋을 웹에 공개 또는 출판하는 주체	M	
Lineage (데이터 이력정보)	데이터셋을 생성하는데 사용된 소스 및 처리 단계에 대한 기록	R	추가
Alternate Identifier (대체 식별자)	등록 중인 자원에 적용된 1차 식별자 이외의 식별자	R	
Metadata of Metadata (메타데이터를 설명한 데이터)	메타데이터를 설명한 데이터	R	추가
Access Right (데이터 접근 권한)	자원의 접근(이용) 권한이나 보안 상태에 대한 정보	R	추가
Contributor (데이터 기여자)	데이터셋의 수집, 생산, 예산지원 등에 관계된 사람이나 기관 등을 기술함	R	
Language (데이터 기술 언어)	정보자원의 기본 언어	O	
Access Restriction (접근 제한)	데이터셋의 접근제한 또는 이용제한 유형	O	
Size (데이터 크기)	파일의 크기(용량)를 단위 없이 값만 기술함	O	
On-line resource (온라인 자원)	데이터셋, 명세, 혹은 단체 프로파일명과 확장된 메타데이터 요소가 획득될 수 있는 온라인 자원에 대한 정보	O	추가
Specification (규격)	데이터 형식 또는 연구 커뮤니티 요구사항 등과 관련된 규격	O	추가
Degree (적합성 정도)	데이터 형식 및 내용에 요구되는 규격 준수 여부	O	추가
Distribution format (데이터 제공 형식)	배포되는 자원(공간정보)의 형식정보	O	추가
Funding Reference (펀딩 참조정보)	연구비 정보로서 연구비 지원기관 정보, 연구비, 프로젝트명, 프로젝트 코드 등을 기술	O	
Audience (이용 대상)	해당 정보자원을 유용하게 사용할 수 있는 이용 대상	O	추가

5. 결 론

많은 학문 분야에서는 해당 분야의 데이터를 효율적으로 관리하기 위하여 메타데이터 개발이 시도되고 있다. 하지만 아직까지 임산공학 분야에서 데이터를 관리하기 위한 메타데이터 개발에 관한 연구는 전무한 실정이다. 따라서 이번 연구에서는 임산공학에서 생산되는 연구 데이터를 관리하기 위한 메타데이터 요소 도출을 연구목적으로 정했다. 메타데이터 요소를 선정하기 위하여 관련된 메타데이터 표준으로 FGDC-CSDGM, ISO 19115, NEFIS, INSPIRE, ANZLIC, TTAK.K0-10.0976 그리고 DataCite 메타데이터 스키마 4.3 등에 대한 내용을 분석하였다.

특히, 메타데이터 상호운용성을 확보하기 위하여 본 연구에서는 메타데이터 간 크로스워크(Crosswalk)를 진행하였다. 진행 결과, 총 28개 항목을 도출하였으며, 도출된 항목의 검증을 위하여 해당 분야 연구자를 대상으로 검증을 실시하였다. 즉, 국립산림과학원 소속 연구자를 대상으로 온라인 설문조사를 시행하였으며 설문조사 결과를 반영하여 분석기준을 수립하고 이를 적용하여 메타데이터 요소를 재선정

하였다.

그 결과 임산공학 분야에서 생산되는 연구 데이터를 기술하기 위한 메타데이터는 총 28개로 도출되었으며 도출된 메타데이터 중 6개를 '필수 항목'으로, 13개를 '권고 항목'으로, 나머지 9개를 '선택 항목'으로 구분하였다. '필수 항목'에는 'Title', 'Publication Year', 'Resource Type', 'Creator', 'Identifier' 그리고 'Publisher' 등이 속한다. '권고 항목'에는 'Subject', 'Date', 'Format', 'Description', 'Geographic Location', 'Rights', 'Version', 'Related Identifier', 'Lineage', 'Alternate Identifier', 'Metadata of Metadata', 'Access Right' 그리고 'Contributor' 등이 포함된다. '선택 항목'에는 'Language', 'Access Restriction', 'Size', 'On-line resource', 'Specification', 'Degree', 'Distribution format', 'Funding Reference' 그리고 'Audience' 등이 해당한다.

이번 연구를 통하여 얻을 수 있는 기대 효과는 해당 연구에서 도출한 메타데이터를 이용하여 임산공학에서 생산된 데이터를 체계적으로 관리할 수 있다는 것이다. 즉, 추후 임산공학 분야의 데이터를 관리하기 위한 시스템 설계 시 데이터의 검색, 공유, 재사용 등 관리를 위한 항목으로 사용될 수 있다는 것이다.

참 고 문 헌

- [1] 고영만, 서태설, 임태훈. 2007. 의미 호환을 위한 메타데이터 매핑 연구. 『정보관리학회지』, 24(4): 223-238.
- [2] 김경민, 김철민, 김태균. 2008. 수치임상도 표준 메타데이터 설계 및 구현. 『한국지리정보학회지』, 11(4): 51-63.

- [3] 김주섭, 김선태, 최상기. 2020. Geoscience 연구데이터 관리를 위한 기능별 세부요소 및 중요도에 관한 연구. 『한국문헌정보학회지』, 54(1): 411-440.
- [4] 남태우, 이승민. 2018. 『정보자원의 기술과 메타데이터』. 서울: 한국도서관협회.
- [5] 노영희. 2012. 『메타데이터 개론』. 서울: 조은글터.
- [6] 박옥남. 2012. PREMIS 기반 보존 메타데이터 요소 개발에 관한 연구. 『한국문헌정보학회지』, 46(2): 83-113.
- [7] 서강현 외. 2017. 풍수해 피해예측지도 연계·활용을 위한 표준 메타데이터 설계. 『한국지리정보학회지』, 20(3): 52-66.
- [8] 심규성. 2016. 침수지도 수집 및 관리를 위한 벡터형 공간정보 표준 기반의 메타데이터 설계. 『한국산학기술학회 논문지』, 17(5): 42-48.
- [9] 예상준, 장호, 김선태. 2019. 한의학 연구 데이터 관리 및 공유를 위한 메타데이터 요소 설계. 『한국문헌정보학회지』, 53(2): 223-246.
- [10] 이미화, 이은주, 노지현. 2020. 연구데이터 관리를 위한 OAK 메타데이터 확장 방안 연구. 『한국도서관·정보학회지』, 51(3): 27-51.
- [11] 전방진. 2011. 활용도 제고를 위한 지적정보의 메타데이터 표준화 적용방안 『한국지적정보학회지』, 13(1): 81-92.
- [12] 한국정보통신기술협회. 2017. 『연구데이터 관리 및 공유를 위한 메타데이터(TTAK.KO-10.0976)』. 성남: 국정보통신기술협회.
- [13] ANZLIC. 2011, ANZLIC Metadata Profile Guidelines Version 1.2. [online] [cited 2019. 7. 2.] <https://www.anzlic.gov.au/sites/default/files/files/anzlicmetadataprofileguidelines_v1-2.pdf>
- [14] CodeMeta. 2019. CodeMeta Overview. [online] [cited 2020. 6. 25.] <<https://codemeta.github.io/developer-guide/>>
- [15] DCC. 2019. Federal Geographic Data Committee Content Standard for Digital Geospatial Metadata. [online] [cited 2020. 7. 1.] <<http://www.dcc.ac.uk/resources/metadata-standards/fgdcccsgm-federal-geographic-data-committee-content-standard-digital-ge>>
- [16] DCC. 2019. ISO 19115. [online] [cited 2020. 7. 2.] <<http://www.dcc.ac.uk/resources/metadata-standards/iso-19115>>
- [17] Desconnets, J. C., Mougnot, I. and Chahdi, H. 2017. "A methodology for effective metadata design in Earth observation. In Developing Metadata Application Profiles." *IGI Global*, 65-97.
- [18] European Commission. 2008. Commission Regulation (EC) No 1205/2008 of 3 December 2008 implementing Directive 2007/2/EC of the European Parliament and of the Council as regards metadata (Text with EEA relevance) Metadata Implementing Rules Part B. [online] [cited

2020. 7. 2.] <<https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX:32008R1205>>
- [19] European Forest Institute. 2019. [online] [cited 2020. 7. 26.] <<https://www.efi.int/>>
- [20] Habermann, T. 2019. "Mapping ISO 19115-1 geographic metadata standards to CodeMeta." *PeerJ Computer Science*, 5: e174.
- [21] Heery, R., and Patel, M. 2000. "Application profiles: mixing and matching metadata schemas." *Ariadne*, (25).
- [22] ISO. [n.d.]. International Organization for Standardization (1 April 2014). "ISO 19115-1:2014 (en)." [online] [cited 2020. 7. 15.] <<https://www.iso.org/standard/53798.html>>
- [23] Li, C. and Sugimoto, S. 2018. "Provenance description of metadata application profiles for long-term maintenance of metadata schemas." *Journal of Documentation*, 74(1): 36-61.
- [24] Malta, M. C. and Baptista, A. A. 2012. "State of the art on methodologies for the development of a metadata application profile." *In Research Conference on Metadata and Semantic Research*, 61-73.
- [25] Morsy, M. et al. 2017. "Design of a metadata framework for environmental models with an example hydrologic application in HydroShare." *Environmental modeling & software*, 93: 13-28.
- [26] NEFIS - Network for a European Forest Information Service. 2019. [online] [cited 2020. 5. 27.] <<https://www.efi.int/projects/nefis-network-european-forest-information-service.>>
- [27] NOAA Environmental Data Management Wiki. 2010. ISO 19115 Core Elements. [online] [cited 2020. 7. 29.] <https://geo-ide.noaa.gov/wiki/index.php?title=ISO_19115_Core_Elements>
- [28] Schuck, A. et al. 2006. "A metadata schema for forest information resources." *The FBMIS Group*, 1: 11-21.
- [29] Schuck, Andreas, 2007. *Towards a European forest information system*, Boston: Brill.
- [30] Wikipedia, 2019. PROV (Provenance). [online] [cited 2020. 6. 25.] <[https://en.wikipedia.org/wiki/PROV_\(Provenance\)](https://en.wikipedia.org/wiki/PROV_(Provenance))>
- [31] Wikipedia, 2019. Directive (European Union). [online] [cited 2020. 7. 2.] <[https://en.wikipedia.org/wiki/Directive_\(European_Union\)](https://en.wikipedia.org/wiki/Directive_(European_Union))>

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

- [1] Ko, Young-Man, Seo, Tae-Sul and Lim, Tae-Hoon. 2007. "A Study on Metadata Mapping

- for Semantic Interoperability.” *Journal of the Korean Society for Information Management*, 24(4): 223-238.
- [2] Kim, Kyoung-Min, Kim, Cheol-Min and Kim, Tae-Kyun. 2008. “Design and Implementation of Standard Metadata for Digital Forest Cover Type Map.” *Journal of the Korean Association of Geographic Information Studies*, 11(4): 51-63.
- [3] Kim, Juseop, Kim, Suntae and Choi, Sangki. 2020. “A Study on Functional Details and Importance of Geoscience Research Data Management.” *Journal of the Korean Society for Library and Information Science*, 54(1): 411-440.
- [6] Park, Ok-Nam. 2012. “A Study on Developing Preservation Metadata Based on PREMIS Focusing on Digital Data in National Library of Korea.” *Journal of the Korean Society for Library and Information Science*, 46(2): 83-113.
- [7] Seo, Kang-Hyeon et al. 2017. “Standard Metadata Design for Linkage and Utilization of Damage Prediction Maps.” *Journal of the Korean Association of Geographic Information Studies*, 20(3): 52-66.
- [8] Sim, Gyu-Seong. 2016. “Metadata Design Based on Vector Type Geospatial Information Standard for the Collection and Management of Inundation Map.” *Journal of the Korea Academia-Industrial cooperation Society*, 17(5): 42-48.
- [9] Yea, Sang-Jun, Jang, Ho and Kim, Suntae. 2019. “Metadata Element Design for Korean Medicine Research Data Management and Re-use.” *Journal of the Korean Society for Library and Information Science*, 53(2): 223-246.
- [10] Lee, Mihwa, Lee, Eun-Ju and Rho, Jee-Hyun. 2020. “A Preliminary Study on Extending OAK Metadata for Research Data.” *Journal of Korean Library and Information Science Society*, 51(3): 27-51.
- [11] Jeon, Bang Jin. 2011. “A Study on the Applying Metaata Standards in Cadastral Information for increasing its Utilization.” *Journal of the Korean Cadastre Information Association*, 13(1): 81-92.

[부록] 임산공학 분야 연구데이터 관리를 위한 메타데이터 요소 및 스키마

(부록-1) 목재제품을 위한 메타데이터 요소와 스키마

ID	메타데이터 요소	표시 상수	빈도	필수 여부	정의, 이용, 값, 예제, 제한 조건
1	Identifier	식별자	1	M	<ul style="list-style-type: none"> • 데이터셋의 식별자 또는 식별값 • 공식적인 식별 체계에 맞는 문자열이나 번호를 사용 • 공식적인 식별 체계에 제한이 없으나, URI(웹주소인 URL을 포함), 디지털 객체 식별기호(DOI), 국가표준 디지털콘텐츠식별체계(UCI) 등이 포함됨 • 데이터셋을 웹에 공개 또는 출판하기 위해 반드시 기술되어야 함
2	Title	제목	1-n	M	<ul style="list-style-type: none"> • 데이터셋의 이름 또는 명칭 • 기술되는 값의 언어 정보가 함께 기술되는 것을 권고함 • 데이터셋을 웹에 공개 또는 출판하기 위해 반드시 기술되어야 함
3	Creator	생산자	1-n	M	<ul style="list-style-type: none"> • 데이터셋을 생성 또는 묶은 사람, 조직 등을 기술함 • 기술되는 값의 언어 정보가 함께 기술되는 것을 권고함 • 데이터셋을 웹에 공개 또는 출판하기 위해 반드시 기술되어야 함 • 활용예시: Tesoro, S. 또는 김선태
4	Publisher	출판자	1	M	<ul style="list-style-type: none"> • 데이터셋을 웹에 공개 또는 출판하는 주체 • 연구자, 기관, 서비스명 등을 기술함 • 데이터셋을 웹에 공개 또는 출판하기 위해 반드시 기술되어야 함
5	Publication Year	출판년도	1	M	<ul style="list-style-type: none"> • 데이터셋이 웹상에서 접근 또는 이용 가능해진 연도 • 데이터셋의 공개 범위가 'Embargoed(이용 지연)'으로 설정되면, 이용 지연이 종료되는 연도를 기술함 • YYYY 형식의 4자리 연도로 기술함 • 데이터셋을 웹에 공개 또는 출판하기 위해 반드시 기술되어야 함 • 활용예시: 2019
6	Date	날짜	0-n	R	<ul style="list-style-type: none"> • 데이터셋이 생애주기 동안 발생한 다양한 이벤트 날짜 • 기술되는 값은 ISO 8601[W3CDTF]에 정의된 방식 중 'YYYY-MM-DD' 형식을 사용할 것을 권고함
7	Language	언어	0-1	O	<ul style="list-style-type: none"> • 데이터셋을 기술한 언어 • Languagecode: ISO 639-3 사용
8	Description	설명	0-n	R	<ul style="list-style-type: none"> • 데이터셋의 내용에 대한 설명 • 기술방법에는 제한이 없으나 주로 맥락 정보를 상세히 기술 • 활용예시: 이 데이터셋은 OOO에서 1시간 주기로 수집된 데이터셋으로서 재사용을 위해서는 OOO 소프트웨어가 필요함
9	Subject	주제	0-n	R	<ul style="list-style-type: none"> • 데이터셋의 내용이 지닌 주제 정보를 기술함 • 활용예시: Human Geography

ID	메타데이터 요소	표시 상수	빈도	필수 여부	정의, 이용, 값, 예제, 제한 조건
10	Contributor	기여자	0-n	R	<ul style="list-style-type: none"> • 데이터셋의 수집, 생산, 예산지원 등에 관계된 사람이나 기관 등을 기술함 • 활용예시: EPSRC - Engineering and Physical Sciences Research Council
11	Rights	저작권	0-1	R	<ul style="list-style-type: none"> • 데이터셋의 저작권 정보를 기술함 • 활용예시: 해당 데이터셋의 저작권은 연구자에게 귀속됨
12	Access Restriction	접근 제한	0-1	O	<ul style="list-style-type: none"> • 데이터셋의 접근제한 또는 이용제한 유형 • 데이터셋 이용에 비용이 필요한지 (feeRequired), 등록이 필요한지 (registration), 무료인지 (free)를 기술 • 활용사례: feeRequired (해당 데이터셋이 유료인 경우)
13	Resource Type	데이터셋 유형	1	M	<ul style="list-style-type: none"> • 데이터셋 유형을 자유롭게 기술 • 인구통계 데이터셋인지, 컨퍼런스 초록데이터 인지 등을 자유롭게 기술 • 활용사례: Dataset/Census Data, Text/Conference Abstract
14	AlternateIdentifier	대체 식별자	0-n	R	<ul style="list-style-type: none"> • 대체 식별자 • 시스템 로컬 식별자 등을 기술함 • Identifier에서 가리키는 동일한 데이터를 가리킴 • 동일한 파일과 동일한 연구 데이터 위치를 가리킴
15	RelatedIdentifier	관련 식별자	0-n	R	<ul style="list-style-type: none"> • 연구 데이터와 관련된 자원 식별자를 기술함 • 글로벌 식별자를 기술해야 함
16	Funding Reference	펀딩 참조정보	0-n	O	<ul style="list-style-type: none"> • 연구비 정보로서 연구비 지원기관 정보, 연구비, 프로젝트명, 프로젝트 코드 등을 기술 • 활용사례: 연구재단 (프로젝트 코드 - GBMF3859.01)
17	Format	파일형식	0-1	R	<ul style="list-style-type: none"> • 파일 형식을 확장자명과 간단한 설명으로 기입함 • 활용예시: Tiff, gif
18	Version	버전	0-1	R	<ul style="list-style-type: none"> • 데이터셋 버전 • 다양한 정보 업데이트를 나타내기 위해 AlternateIdentifier 및 RelatedIdentifier와 함께 사용할 수 있음 • 버전의 특성 및 파일/레코드 범위를 나타내기 위해 Description과 함께 사용할 수 있음
19	Size	파일크기	0-1	O	<ul style="list-style-type: none"> • 파일의 크기(용량)를 단위 없이 값만 기술함 • 바이트 단위로 입력하는 것을 권고함 • 활용예시: 2,500,000
20	Geographic Location	위치 정보	0-n	R	<ul style="list-style-type: none"> • geoLocationPoint(점경도, 점위도), geoLocationBox(서쪽경계경도, 동쪽경계경도, 남쪽경계위도, 북쪽경계위도), geoLocationPlace(지역 이름이나 지역에 대한 설명), geoLocationPolygon(다각형지역)를 선택적으로 기술하거나 자유로운 텍스트로 기술함
21	Lineage	데이터셋 이력	0-1	R	<ul style="list-style-type: none"> • 연혁에 대한 정보의 부족이나 범위에 의해 상세된 데이터를 만드는데 사용된 이벤트나 원시자료에 대한 정보 • 데이터셋을 구성하는데 사용된 이벤트 및 소스 데이터에 대한 정보 또는 정보 부족에 대한 설명 • 정보 또는 정보 부족에 대한 설명을 자유롭게 기술(free text)

ID	메타데이터 요소	표시 상수	빈도	필수 여부	정의, 이용, 값, 예제, 제한 조건
22	Distribution format	배포 형식	0-1	O	<ul style="list-style-type: none"> • 자원(공간정보)을 얻기 위해 필요한 배포자와 기술적인 옵션들에 대한 정보 • 배포되는 자원(공간정보)의 형식 정보 예) DXF, ARC/INFO, Shape 등
23	On-line resource	온라인 정보자원	0-n	O	<ul style="list-style-type: none"> • 데이터셋, 명세, 혹은 단체 프로파일명과 확장된 메타데이터 요소가 획득될 수 있는 온라인 자원에 대한 정보 • URL 등과 같은 자원(공간정보)을 온라인 상으로 획득할 수 있는 자료의 위치 • 문자열로 표시(URL Scheme: IETF RFC1738, IETF RFC2056)
24	Metadata of Metadata	메타데이터의 데이터	0-1	R	<ul style="list-style-type: none"> • 메타데이터의 메타데이터 • 메타데이터 파일 식별자, 메타데이터 기술 언어, 메타데이터 문자열 셋, 메타데이터 연락담당자, 메타데이터 생성날짜, 메타데이터 표준명, 메타데이터 버전 등을 선택적으로 기술
25	Audience	이용대상	0-1	O	<ul style="list-style-type: none"> • Dublin Core의 선택 요소 • 해당 정보자원이 의도하거나 유용하게 사용할 수 있는 엔티티 클래스 • 이용대상을 자유롭게 기술(free text)
26	Access Right	접근 권한	0-1	R	<ul style="list-style-type: none"> • 자원의 접근(이용) 권한이나 보안 상태에 대한 정보 • 자유롭게 기술(free text)
27	Specification	규격	0-1	O	<ul style="list-style-type: none"> • 데이터 형식 또는 연구 커뮤니티 요구사항 등과 관련된 규격 • 해당 규격의 제목과 참조일(발행일, 최종 개정일 또는 생성일) 포함하여 기술
28	Degree	적합성 정도	0-1	O	<ul style="list-style-type: none"> • 데이터 형식 및 내용에 요구되는 규격 준수 여부 • 규격 통과 여부를 수치로 입력('0' 또는 '1') 예) 0: False, 1: Pass