

국가과학기술표준분류체계 용어 관리를 위한 SKOS 기반 메타데이터 요소 개발 연구*

A Study on Development of SKOS-based Metadata Elements for Managing Keywords in the National Science and Technology Standard Classification System

송 민 선 (Min Sun Song)**

박 진 호 (Jin Ho Park)***

초 록

국가과학기술표준분류는 과학기술 관련 정보, 인력, 연구개발사업 등을 효율적으로 관리하기 위한 목적으로 제정 및 운영되고 있으며 개정주기는 5년이다. 2022년은 차기 개정 절차의 첫 해로 현재의 대, 중, 소분류체계 중 소분류체계를 기술키워드화 하는 것이 주 목적이다. 이는 현재의 경직된 구조로 인해 발생하는 유관 분류체계와의 연계 어려움과 최신 용어에 대한 미반영 문제를 해결하기 위한 것이다. 본 연구에서는 이 문제 해결을 위해 기존의 분류체계 관리를 용어관리체계로 변화시켜 용어의 품질과 활용성을 높일 수 있는 방법을 제안하였다. 이를 위해 표준용어관리체계인 SKOS와 ISO/IEC 11179 표준을 기본 모델로 설정하였다. 또 해외 과학기술용어집에서 활용하고 있는 용어관리 메타데이터 표준을 조사하여 현 국가과학기술표준분류체계와 비교한 후 용어관리관점에서 즉시 활용할 수 있는 메타데이터들을 도출하였다. 그 결과 현 관리체계에서 즉시 변형하여 적용할 수 있는 11개 표준 요소를 발굴 제안하였으며, 차후 분류체계 개정 작업을 거친 후 적용할 수 있는 5개 요소를 발굴하여 제안하였다.

ABSTRACT

The National Science and Technology Standard Classification System is established and operated for the purpose of efficiently managing science and technology related information, manpower, and R&D projects. The revision cycle for the classification system is five years, and 2022 is the first year of the next revision procedures. The main purpose of the next revision is to turn the third level categories into the keywords in the current classification system. It is to solve the problems about not reflecting the latest terms, and the difficulty in linking with the relevant other classifications caused by the rigid structure of the current classification system. In this study, the method was proposed by changing the existing classification system into the term management system as to improve the quality and usability of keywords related with the current third level categories. For this method, SKOS, the international standard terminology management system, and ISO/IEC 11179 standards were offered as basic models. In addition, the related metadata standards used in overseas scientific and technological glossaries were investigated, and compared with the current National Science and Technology Standard Classification System. And then essential metadata elements from the terminology management as point of view was derived. As a result, 11 standard metadata elements that can be immediately modified and applied in the current system were recommended, and five elements that can be applied after the revision of the classification system were offered.

키워드: 국가과학기술표준분류체계, SKOS, 메타데이터 요소, ISO/IEC 11179, 용어관리

National Science and Technology Standard Classification System, SKOS, Metadata Elements, ISO/IEC 11179, Terminology Management

* 이 논문은 2021년 한국과학기술기획평가원의 국가과학기술표준분류 소분류 기술키워드화 운영방안 설계 연구 용역의 일부를 수정·보완한 것임.

** 대림대학교 도서관미디어정보과 조교수(songery@daelim.ac.kr) (제1저자)

*** 한성대학교 크리에이티브 인문학부 도서관정보문화트랙 조교수(jhp@hansung.ac.kr) (교신저자)

논문접수일자 : 2021년 11월 21일 논문심사일자 : 2021년 11월 22일 게재확정일자 : 2021년 12월 2일
한국비블리아학회지, 32(4) : 67-88, 2021. <http://dx.doi.org/10.14699/kbiblia.2021.32.4.067>

© Copyright © 2021 Korean Biblia Society for Library and Information Science

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

1. 서론

국가과학기술표준분류체계(이하, 표준분류체계)는 과학기술 관련 정보, 인력, 연구개발사업 등을 효율적으로 관리하기 위한 목적으로, 『과학기술기본법』 제27조 및 시행령 제41조에 근거해 2002년 최초 제정되었다. 이후 3년마다 수정·보완되어 오다 2012년에 5개년 개정 주기로 프로세스가 변경되었다. 표준분류체계는 과학기술정보나 인력, 연구개발 사업 관리 목적 외에도 정부 및 민간 영역의 투자 현황 분석이나 국내외 R&D 기초 통계 분류 기준으로써도 활용되고 있으며, 정부의 관련 사업 예산 조정 및 배분 등 다양한 분야에서 활용되고 있다(최문정 외, 2019).

가장 최근에 활용하고 있는 표준분류체계는 2013년부터 2017년까지 단계별 개정 작업을 추진해 2017년 확정된 분류 개정(안)으로, 차기 표준분류체계 개정은 2022년으로 계획하고 있다. 5년마다 추진되는 개정 프로세스는 수요 발굴부터 개정까지 5단계의 절차를 거치는데, 2022년 차기 개정 작업을 앞둔 2021년 현재에는 분류체계 개정 근거 및 영향력에 대한 모니터링 작업을 수행하는 3단계 절차까지 진행되었다. 이에 앞서 2018년에는 차기 분류체계의 개정을 위한 수요 도출을 위해 개정 수요 조사를 실시(1단계)하였고, 2019년에는 개정 수요 내용을 대상으로 적합성 평가를 통해 개정 타당성 평가대상을 선정(2단계)하였다. 현재 진행되는 3단계 모니터링 결과가 도출되면 이 내용을 토대로 2022년에는 개정 타당성 평가 작업 이후(4단계) 최종적으로 정규 분류 개정(5단계)을 추진하게 된다.

2017년 현행 분류체계 개정 작업 후 표준분류체계 제도 개선에 대해 진행했던 설문 결과, 현행 분류체계 활용에 있어 개인연구자들의 경우에는 급변하는 신기술 분야에 대한 업데이트가 늦어 연구 분야 분류 작업 시 필요한 항목이 누락된다는 문제를, 전문 기관의 경우에는 자체적인 분류체계와 상이한 내용의 불편함에 대한 문제를 제기하였다. 이러한 문제를 해결하기 위해 현재의 '4계층 구조(연구분야>대>중>소)'를 '3계층 구조(연구분야>대>중)'으로 단순화하고 기존 소분류를 기술 키워드로 대체·활용하는 방안을 추진 중이다(과학기술정보통신부, 2019). 따라서 2022년에 앞두고 있는 차기 표준분류체계 개정은 분류구조를 단순화하고 신기술 등에 대한 신속한 분류 반영을 위해 유관 분류체계와 연계를 강화함으로써 활용도를 제고하는 것을 목적으로 한다. 이를 위해 기존 2022년 개정 표준분류체계의 소분류를 기술키워드로 대체·적용하기 위한 구체적인 실행 방안 수립이 필요하다.

본 연구의 목적은 2022년부터 시행할 소분류의 기술키워드화 수행 시 이 기술키워드를 표준용어관리체계에 맞추어 변환하고 향후 관리할 수 있도록 표준 메타데이터를 요소를 제안하는 것이다. 이는 향후 기술키워드의 상호운용성을 확보하여 활용성을 높이고 용어자체의 품질을 높이기 위한 것이다.

이를 위해 먼저 국제적으로 용어관리와 관련해 가장 많이 활용되고 있는 SKOS(Simple Knowledge Organization System)와 ISO/IEC 11179 표준을 검토하였다. 이는 향후 소분류 체계를 기술키워드 방식으로 대체 운영하게 되었을 때 지속적으로 추가·갱신될 신규 용어들을

효과적으로 관리하는 데 필요한 메타데이터 요소들을 확인하고자 하는 것이다. 또한 과학기술 분야와 관련된 전문용어들을 관리하는 해외 주요 용어관리체계 사례들을 검토하고, 각 사례들에서 활용하고 있는 메타데이터 요소들을 분석해 향후 표준분류체계 기술키워드 관리 운영에 실질적으로 필요한 메타데이터를 제안하였다.

2. 기술키워드 용어 관리를 위한 메타데이터 관련 표준

2.1 SKOS(Simple Knowledge Organization System)

SKOS는 W3C에서 개발한 시맨틱 웹 기반의 지식 구조화 프레임워크 표준으로, RDF(Resource Description Framework)를 기반으로 한다. 시소러스, 분류체계, 택소노미, 색인 목록을 비롯한 다양한 유형의 구조적인 어휘들을 표현할 수 있도록 디자인되어 있어 일반적인 구조적 어휘 요소들을 표현하기에 적합한 클래스와 속성을 정의하고 있다(〈표 1〉 참조).

SKOS는 다양한 형태의 지식 자원을 표현하는 데 있어 다음과 같은 특징들을 갖는다.

첫째, SKOS는 skos:altLabel, skos:prefLabel, skos:hiddenLabel, skos:prefSymbol, skos:altSymbol 등의 다양한 레이블링 방법을 제시하고 있어, 대상 자원의 다국어 표현, 기호, 상징 등을 모두 표현할 수 있다는 장점을 갖는다.

둘째, 개념 정의(skos:definition), 활용 범위 명시(skos:scopeNote), 과거 이력 확인(skos:historyNote), 사용 사례(skos:example) 등 기

술 대상 자원을 효과적으로 관리할 수 있는 문서화 속성을 제공하고 있다.

셋째, SKOS는 시소러스에서 제공되는 기본적인 의미 관계인 BT(상위어, skos:broader), NT(하위어, skos:narrower), RT(관련어, skos:related) 등에 대한 표현 방식을 제공하고 있어 기술된 개념 간의 관계를 정의할 수 있다.

넷째, skos:subject, skos:isSubjectOf, skos:PrimarySubject, skos:isPrimarySubjectOf 등의 주제 색인을 위한 속성을 가지고 있어, 대상 자원의 개념이 어떤 주제와 관련되는지를 나타낼 수 있다.

마지막으로, 구체적으로 기술된 개념들을 기준에 따라 그룹화해 표현할 수 있는 기능도 제공한다. 즉, buffalo milk, cow milk, goat milk, sheep milk의 개념들을 milk by source animal이라는 그룹으로 표현할 수 있다.

이처럼 SKOS는 사용하는 기관별 용도나 목적에 따라 확장이 용이하기 때문에 원래 지향하는 목적을 유지하면서도 각각의 개별 도메인에 특화된 특성들을 표현할 수 있어, LOD(Linked Open Data)로 공개·공유되는 데이터 집합들은 SKOS 기반으로 구축되는 경우가 많다.

2.2 ISO/IEC 11179

ISO/IEC 11179는 ISO/IEC JTC1/SC32(Data Management and Interchange)에서 관리하는 표준이다. 이는 메타데이터를 공유하기 위해 필요한 메타데이터 레지스트리(Metadata Registry, 이하 MDR)에 관한 표준이다. 이 표준에서는 메타데이터 요소의 등록, 관리, 유지, 관리에 필요한 요소와 체계를 포괄적으로 제시

〈표 1〉 SKOS의 클래스와 속성(Miles & Brickley, 2005)

구분	구성요소	정의
Class	skos:CollectableProperty	skos:Collection과 함께 사용할 수 있는 속성을 적용
	skos:Collection	의미 있는 개념 모음
	skos:Concept	추상적인 아이디어 또는 개념 또는 생각의 단위
	skos:ConceptScheme	선택된 개념 간 의미관계 설명을 포함하는 개념 체계
	skos:OrderedCollection	그룹 및 순서가 적용된 개념 모음
Property	skos:altLabel	기술 대상 자원에 대한 대체 용어를 기술(약어, 철자 변형, 불규칙 단복수 등의 용어를 포함하며 철자가 틀린 용어는 hiddenLabel에 기술)
	skos:prefLabel	특정 언어로 된 자원에 대한 기본 레이블
	skos:hiddenLabel	대상 자원의 시각적 표시는 숨겨두되, 검색 내용에는 포함시켜야 하는 내용을 기술
	skos:prefSymbol	기술 대상 자원에 대한 기본 기호를 기술
	skos:altSymbol	기술 대상 자원에 대한 대체 기호를 기술
	skos:definition	대상 개념의 의미에 대한 설명
	skos:scopeNote	개념의 의미를 명확히 하는 데 도움이 되는 메모
	skos:example	개념 사용 사례 기술
	skos:historyNote	개념의 과거 상태에 대한 용도 및 의미를 메모
	skos:editorialNote	대상 어휘 편집자 및 관리자를 위한 메모
	skos:changeNote	개념 수정에 대한 참고 사항을 기술
	skos:Note	일반적인 메모
	skos:semanticRelation	의미와 관련된 개념(이 속성은 직접 사용하지 않으며, 개념 간 의미 관계를 나타내는 모든 속성의 상위 속성으로 사용함)
	skos:broader	상위 개념(상위어)
	skos:narrower	하위 개념(하위어)
	skos:related	관련 개념(관련어)
	skos:subject	자원의 주제인 개념
	skos:isSubjectOf	개념이 주제인 자원
	skos:primarySubject	문서의 주요 주제인 개념
	skos:isPrimarySubjectOf	대상 개념이 기본 주제인 문서
	skos:member	Collection의 구성요소
	skos:memberList	정리된 Collection의 구성요소를 포함하는 RDF 리스트
	skos:subjectIndicator	개념에 대한 주제 표시자(Indicator)
	skos:symbol	대상 자원에 대한 기호 레이블인 이미지(rdfs:label과 유사하나 문자가 아닌 검색 가능한 표현이 있는 이미지)
	skos:inScheme	대상 개념이 포함된 개념 체계(특정 개념은 두 가지 이상의 개념 체계에 포함될 수 있음)
	skos:hasTopConcept	개념 체계(Concept Scheme)의 최상위 개념

하고 있다. MDR은 메타데이터 요소의 생성, 등록, 관리를 지원하는 시스템으로 이를 통해서 시스템 간에는 물론 조직 간의 메타데이터 정보 공유가 가능하다. 또 이용자들이 메타데이터 요소가 갖는 의미와 표식을 이해하며, 나아가서는

메타데이터의 수집과 서비스에 있어서 일관된 모델을 제시할 수 있다. 이 표준은 ‘메타데이터’를 ‘데이터’로 보고 여기에 필요한 절차, 기술적 요건, 운영방식 등을 상세하게 기술하고 있어, 과학기술분류에 활용하는 용어, 개념 관리 개선

에 응용하여 활용할 수 있다. ISO/IEC 11179는 총 7개 부분으로 구성되며 <표 2>와 같다.

Part1은 동 표준을 구성하는 각 부분들을 이해하고 연관 지을 수 있는 설명을 포함하고 있다. Part2는 MDR 등록항목을 분류하여 Part3의 정보를 보완해준다. Part3은 데이터 관리를 위한 메타 모델을 제시한다. Part4는 명확한 의미 제공이 가능하도록 데이터를 어떻게 정의해야 하는지에 대한 지침을 정의한다. Part5는 요소에 대한 명명, 식별 규칙으로 세부적으로는 지침을 포함하고 있다. Part6은 메타데이터 요소

등록 관리활동에 필요한 역할과 책임사항, 데이터가 생성되고 유지되는 활동, 데이터의 상태 등에 대한 정보를 포함하고 있다. 마지막으로 Part7은 비교적 최근에 만들어진 것으로 데이터 셋이라는 새로운 정보자원이 대상이다. 본 연구에서는 기술키워드의 용어관리 측면은 SKOS 표준을 활용하고 용어의 상태정보관리는 ISO/IEC 11179 표준을 대상으로 하였다. 용어의 상태 관련 정보는 Part6에서 확인할 수 있다. 크게 동적상태와 정적상태로 구분하여 용어의 상태를 표현하는데 이를 정리하면 <표 3>과 같다.

<표 2> ISO/IEC 11179 표준의 구성

구분	명칭
Part1	Framework: 프레임워크
Part2	Classification: 데이터 요소 분류
Part3	Registry Metamodel and Basic Attributes: 레지스트리 메타 모델과 메타데이터의 기본 속성
Part4	Formulation of Data Definitions: 데이터의 정의
Part5	Naming and Identification: 명명과 식별
Part6	Registration: 등록
Part7	Metamodel for data set registration: 데이터셋 등록을 위한 메타 모델

<표 3> ISO/IEC 11179의 등록 데이터 상태별 표현

상태	설명	구분
Preferred Standard	이용자를 위한 최적의 품질	동적상태
Standard	이용자를 위한 높은 품질	동적상태
Qualified	필수 메타데이터 요소 충족, 품질 요구사항에 부합	동적상태
Recorded	필수 메타데이터 요소 충족	동적상태
Candidate	등록 상태 승격 제안 중	동적상태
Incomplete	등록 요청자가 제출한 관리항목	동적상태
Retired	더 이상 사용되지 않고, 사용되어서는 안됨	정적상태
Superseded	사용을 권장하지 않으며, 새로운 관리항목을 사용할 것을 권장	정적상태
Historical	과거에 외부에서 사용됨	정적상태
Standardized Elsewhere	다른 레지스트리 단체에서 표준화 됨	정적상태
Legacy	특정 관리항목이 현재 로컬 영역에 존재하며 애플리케이션에 사용되나, 관련 정보가 거의 없음	정적상태
Application	특정 관리 항목이 현재 로컬에 존재하며 애플리케이션에 사용되나 논리적 수준으로 기술되어 있지 않음	정적상태

3. 과학기술 분야의 전문 용어 관리 체계 사례별 용어관리 메타데이터 요소

현재 표준분류체계의 소분류를 기술 키워드 방식으로 대체 운영하는 문제가 대두된 근본적인 이유는 현행 분류체계가 텍소노미 형태의 경직된 계층형(hierarchical) 구조로 되어 있어, 급변하는 과학기술 분야의 새로운 개념이나 신조어 등을 신속히 반영하기 어려운 전문가 중심의 닫힌 업데이트 방식을 유지하고 있다는 점에서 기인한다.

이러한 문제를 개선하기 위해 앞으로 표준분류체계가 지향해야 할 방향 중 하나는 국내외 유관 기관에서 운영·관리되고 있는 전문용어관리 체계(분류시스템 및 통제어휘집 등)와 연동해 지속적으로 관련 데이터를 입수해 업데이트에 반영하는 기능을 개발하는 것이다. 특히 분류체계에 반영할 용어 업데이트의 경우, 계속 발전하는 과학기술 및 융복합 분야에 대한 전문 지식이 필요하기에 몇 명의 인력에 의한 운영만으로는 한계가 있을 수 있다. 즉 다양한 과학기술 관련 용어집과의 상호운용이 가능한 형태를 유지하고 정비하는 것이 필요하다. 이에 본 장에서는 과학기술 분야와 관련된 해외 주요 용어관리체계 사례들을 검토하고, 각각에서 활용 중인 메타데이터 요소들을 살펴보고자 한다.

3.1 UNESCO nomenclature for fields of science and technology(과학기술 분야의 UNESCO 명명법)

과학기술 분야의 UNESCO 명명법(이하, UNESCO 명명법)은 1973년과 1974년 UNESCO

의 과학기술정책 및 통계국(Division of Science Policy and Statistics for Science and Technology)에서 제안하고 과학 자문위원회(Scientific Advisory Committee)에서 채택한 분류시스템으로, 과학기술 분야 연구 프로젝트와 학위논문의 지식 관리 및 분류를 위해 만들어졌다.

UNESCO 명명법은 크게 Fields(필드)-Disciplines(분야)-Subdisciplines(하위분야)의 3단계 계층구조로 구성되는데, UNESCO 명명법 분류 시스템 상의 여러 범주들은 SKOS에서 단일 개념 체계(skos:Conceptscheme)로 표현하고 있으며, UNESCO 명명법 체계에서 용어 관리에 활용하고 있는 메타데이터 요소는 <표 4>와 같다.

<표 4> UNESCO 명명법에서 활용 중인 메타데이터 요소

구분	메타데이터 요소
SKOS 활용 요소	• skos:broader • skos:hasTopConcept • skos:inScheme • skos:narrower • skos:notation • skos:prefLabel • skos:topConceptOf
기타 추가 관리 요소	없음

※ 관련 링크: <https://skos.um.es/unesco6/unesco6.ttl>

3.2 UNESCO Thesaurus(UNESCO 시소러스)

UNESCO 시소러스는 교육, 문화, 자연과학, 사회 및 인문과학, 커뮤니케이션 및 정보 분야의 주제 분석과 문서 및 출판물 검색에 사용되는 통제어휘 목록이다. Education(교육), Science(과학), Culture(문화), Social and human sciences(사회과학 및 인문학), Information and communication(정보 및 커뮤니케이션),

Politics, law and economics(정치, 법률, 경제학), Countries and country groupings(국가그룹)의 7가지의 주요 주제 영역으로 구성되어 있으며, 각 주제별로 빠르게 개관할 수 있도록 하위에 마이크로 시소러스(microthesauri)로 분류되는 체계를 갖는다. UNESCO 시소러스에서 용어 관리에 활용하고 있는 메타데이터 요소는 <표 5>와 같다.

<표 5> UNESCO 시소러스에서 활용 중인 메타데이터 요소

구분	메타데이터 요소
SKOS 활용 요소	• skos:Concept • skos:ConceptScheme • skos:altLabel • skos:broader • skos:hasTopConcept • skos:inScheme • skos:member • skos:narrower • skos:notation • skos:prefLabel • skos:related • skos:scopeNote • skos:topConceptOf • skos:hasMicroThesaurus
기타 추가 관리 요소	• uneskos:memberOf • uneskos:mainConceptOf • uneskos:hasMainConcept • uneskos:contains

※ 관련 링크: <https://skos.um.es/unescothes/stats.php>

3.3 Realfagstermer(과학주제용어)

Realfagstermer는 오슬로 대학의 과학도서관(University of Oslo Science Library)에서 베르겐 대학의 과학도서관(University of Bergen Science Library)과의 협력 하에 공동으로 유지 관리하고 있는 과학 분야의 통제어휘집으로, 대체로 물리학, 수학, 컴퓨터 과학 분야 관련 개념들을 포함한다. 2020년 현재 대략 11,000개의 개

념(concepts)과 15,000개의 복합표목(compound headings, 복합주제어)으로 구성되어 있으며, 모든 사람들이 자유롭게 이용할 수 있도록 공개 데이터로 제공된다.

Realfagstermer의 각각의 개념들은 주제(topical), 지리(geographical), 연대(chronological), 장르 및 형식(genre/form)의 4가지 유형으로 구성된다. 각 개념들에는 노르웨이어(Norwegian), 뉘노르스크어(Nynorsk), 영어(English)로 표현되는 고유 식별자와 우선어(preferred term)가 포함되며, 각 개념에는 동의어(synonyms)나 두문자어(acronyms) 같은 대체어(alternative terms)와 학명(scientific names), 관련 참조(도보라(see also) 참조)와 주석(notes, 내부 참조용), 정의(definitions, 최종 이용자용), 그리고 매핑(mapping) 정보가 포함될 수 있다. Realfagstermer에서 관리하는 요소는 <표 6>과 같다.

<표 6> Realfagstermer에서 활용 중인 메타데이터 요소

구분	메타데이터 요소
SKOS 활용 요소	• skos:Concept • skos:ConceptScheme • skos:altLabel • skos:broader • skos:definition • skos:editorialNote • skos:historyNote • skos:inScheme • skos:prefLabel • skos:related • skos:closeMatch • skos:mappingRelation • skos:broadMatch
기타 추가 관리 요소	• dc:modified • dc:identifier • dc:created • ubo:libCode

※ 관련 링크: <https://data.ub.uio.no/skosmos/realfagstermer/en/>

3.4 PLOS Thesaurus(PLOS 시소러스)

PLOS(Public Library of Science)는 과학 분야의 대표적인 오픈 액세스 저널 출판단체이다. PLOS 시소러스는 PLOS에서 다루고 있는 주제 영역들과 관련된 용어를 계층 구조화해 관리하고 있는 통제어휘집으로, MAI(Machine Aided Indexing) 방식을 통해 주제 용어들을 추출해 시소러스에 반영하고 있다. 즉, PLOS 기사 내에서 텍스트 문자열을 식별해 기사 내 출현 빈도를 기준으로 상위 8개 용어를 자동 선택한 후, 지정된 규칙에 따라 선택된 용어와 관련된 맥락(context)을 고려해 용어를 명확히 해 채택하는 방식을 사용하고 있다.

PLOS 시소러스는 PLOS에서 기사 검색을 할 때 활용되는데, 각 PLOS 기사의 주제 영역 패널에는 사용 가능한 1만 여개 용어 중 해당

기사와 관련해 선택된 8개 용어가 표시된다. 이 용자가 해당 기사의 특정 주제 영역 용어를 클릭하면, 그 용어와 관련된 모든 기사 목록을 볼 수 있고, 각 기사의 XML을 다운로드하면 해당 기사 페이지에 연결된 8개 주제 영역 용어와 함께 해당 용어와 연계된 더 광범위한 용어까지 확인할 수 있다. 2020년 1월 현재, PLOS 시소러스는 12단계의 계층구조를 가지며, 1만 8천 여개의 용어를 포함하고 있다. PLOS 시소러스 관리 메타데이터 요소는 <표 7>과 같다.

3.5 OSTI Semantic Thesaurus(OSTI 시맨틱 시소러스)

OSTI(Office of Scientific and Technical Information)는 미국 에너지부(Department of Energy) 산하의 과학기술정보국이다. OSTI 시맨틱 시소러스는 과학 및 기술 용어와 그 용어 관계에 대한 참고 자료로, OSTI.gov 검색에서 키워드와 개념 매핑 구현에 활용되어 최종 이용자가 검색한 용어와 관련된 개념들을 구체적으로 탐색할 수 있도록 도와준다. OSTI 시맨틱 시소러스는 INIS/ETDE 시소러스(국제원자력 정보시스템(International Nuclear Information System, INIS) 및 에너지 기술 데이터 교환(Energy Technology Data Exchange, ETDT)의 주제 범위를 포함하는 시소러스)에서 데이터와 구조를 상속 받았으며, 그 이후 Wikidata 등에서 제공되는 과학적 개념과 수동적 큐레이션(manual curation) 작업을 통해 확장되었다. OSTI에서 활용하는 메타데이터 요소는 <표 8>과 같다.

<표 7> PLOS 시소러스에서 활용 중인 메타데이터 요소

구분	메타데이터 요소
SKOS 활용 요소	• skos:Concept • skos:ConceptScheme • skos:altLabel • skos:broader • skos:changeNote • skos:editorialNote • skos:hasTopConcept • skos:narrower • skos:prefLabel • skos:related • skos:scopeNote
기타 추가 관리 요소	• foaf:homepage • status • Previous_Classification

※ 관련 링크: <https://raw.githubusercontent.com/PLOS/plos-thesaurus/master/plothes.2020-1.SKOS.xml>

〈표 8〉 OSTI 시맨틱 시소러스에서 활용 중인 메타데이터 요소

구분	메타데이터 요소
SKOS 활용 요소	• skos:Concept • skos:ConceptScheme • skos:broader • skos:definition • skos:hasTopConcept • skos:inScheme • skos:narrower • skos:prefLabel • skos:related • skos:scopeNote • skos:topConceptOf
기타 추가 관리 요소	없음

※ 관련 링크: <https://www.osti.gov/ai/research/dataset-semantic-thesaurus>

3.6 EuroVoc

EuroVoc은 EU의 활동과 관련된 다국어, 다학문 영역을 포괄하는 시소러스로, 23개의 EU 가입국 주요 언어 및 3개의 가입 후보 국가 언어를 아우르며, EU 의회 및 출판국(Publications Office)을 비롯해 유럽 내 각국의 부서 및 관련 기관, 조직들에서 사용된다. EU 및 EU 소속 국가들의 관점을 모두 반영한 의회 활동 등에 대한 분야를 다루기 때문에, EuroVoc에서 다루는 주요 도메인은 정치, 국제관계, 법률, 경제, 무역, 금융, 사회문제, 교육 및 커뮤니케이션, 과학, 비즈니스, 고용 및 근무 조건, 환경, 운송, 농업, 임업, 수산업, 농식품 등 다양한 분야를 포함한다. EuroVoc의 메타데이터 요소는 〈표 9〉와 같다.

〈표 9〉 EuroVoc에서 활용 중인 메타데이터 요소

구분	메타데이터 요소
SKOS 활용 요소	• skos:Concept • skos:altLabel • skos:broader • skos:definition • skos:editorialNote • skos:hasTopConcept • skos:inScheme • skos:narrower • skos:notation • skos:prefLabel • skos:related • skos:scopeNote • skos:topConceptOf
기타 추가 관리 요소	• dc:identifier • dc:created • owl:versionInfo • status • xlNotation • xlScopeNote

※ 관련 링크: <https://op.europa.eu/en/web/eu-vocabularies/dataset/-/resource?uri=http://publications.europa.eu/resource/dataset/eurovoc>

4. 표준분류체계 용어 관리를 위한 메타데이터 요소 제안

4.1 해외 용어관리체계에서 활용 중인 메타데이터 요소 분석

3장에 소개된 해외 주요 용어관리체계들에서 활용 중인 메타데이터 요소들을 살펴보면 대체로 SKOS 요소를 주로 적용해 관리하고 있으며, 기타 각 관리체계 내에서 추가로 필요하다고 판단되는 요소들에 대해서는 더블린 코어 요소를 일부 사용하거나, 자체적으로 필요

한 요소들을 정의해 사용하고 있다(〈표 10〉 참조). 더블린 코어 요소의 경우는 해당 용어를 만든 날짜 및 수정한 날짜를 표기하기 위한 created, modified가 대표적이며, 용어별로 부여되는 코드(식별자)를 위한 identifier 요소를

많이 사용한다. 이 외에도 적용 범위와 용어 상태, 용어 상태가 변경되어 이전 분류가 존재하는 경우 이를 표현하기 위해 필요한 요소들을 자체적으로 정의해 사용하고 있음을 확인할 수 있다.

〈표 10〉 해외 주요 용어관리체계에서 활용 중인 메타데이터 요소 매핑

구분	UNESCO 명명법	UNESCO 시소러스	Realfagstermer	PLOS 시소러스	OSTI 시맨틱 시소러스	EuroVoc
SKOS 활용 요소	skos:topConceptOf	skos:topConceptOf			skos:topConceptOf	skos:topConceptOf
	skos:prefLabel	skos:prefLabel	skos:prefLabel	skos:prefLabel	skos:prefLabel	skos:prefLabel
	skos:broader	skos:broader	skos:broader	skos:broader	skos:broader	skos:broader
	skos:hasTopConcept	skos:hasTopConcept		skos:hasTopConcept	skos:hasTopConcept	skos:hasTopConcept
	skos:inScheme	skos:inScheme	skos:inScheme		skos:inScheme	skos:inScheme
	skos:narrower	skos:narrower		skos:narrower	skos:narrower	skos:narrower
	skos:notation	skos:notation				skos:notation
		skos:Concept	skos:Concept	skos:Concept	skos:Concept	skos:Concept
		skos:ConceptScheme	skos:ConceptScheme	skos:ConceptScheme	skos:ConceptScheme	
		skos:altLabel	skos:altLabel	skos:altLabel		skos:altLabel
			skos:definition		skos:definition	skos:definition
		skos:related	skos:related	skos:related	skos:related	skos:related
			skos:historyNote			
			skos:editorialNote	skos:editorialNote		skos:editorialNote
		skos:scopeNote		skos:scopeNote	skos:scopeNote	skos:scopeNote
			skos:changeNote			
		skos:member				
			skos:closeMatch			
기타 추가 관리 요소			dc:created			dc:created
			dc:modified			
			dc:identifier			dc:identifier
			ubo:libCode			
		uneskos:memberOf				
				foaf:homepage		
				status		fstatus
				Previous_Classification		
						owl:versionInfo
						xiScopeNote
						xiNotation
		uneskos:mainConceptOf				
		uneskos:hasMainConcept				
		uneskos:contains				

각 용어관리 체계별로 사용되는 요소들은 용어 관리의 목적에 따라 각각 다르기 때문에, 단순히 활용하고 있는 메타데이터 요소 수가 많다고 해서 용어 관리를 잘하고 있으며, 관리 용어의 품질이 높다고 판단할 수는 없다. 즉, 각 체계별로 활용 중인 메타데이터 요소들은 용어의 활용 목적과 관리 주체의 용어 생성 및 관리 정책에 관련된 사항들을 반영한다고 할 수 있다. 단, 많은 요소를 활용할수록 각각의 관리 체계에서 제공되는 용어들을 다양한 목적으로 활용하고자 하는 이용자들의 요구사항을 충족시킬 가능성은 높다 하겠다.

〈표 10〉의 6개 용어관리 체계들 중 과반수에 해당하는 4개 이상에서 공통으로 사용하고 있는 메타데이터 요소들을 핵심요소로 다시 추려 보면 〈표 11〉과 같이 정리해 볼 수 있다. 각각의 체계들에서 공통적으로 많이 사용하고 있는 요소들은 대체로 관리 대상 용어 및 관련 개념의 명칭 표현, 상/하위 및 관련 관계 용어 연결, 전체 용어집의 명칭, 그리고 용어 관리에 대한 참고사항들을 기록하기 위한 요소들로 구성된다.

〈표 11〉 해외 주요 용어관리체계에서 많이 활용하는 공통 핵심 요소

공통 핵심 요소	활용 체계
skos:Concept	• UNESCO 시소러스 • Realfagstermer • PLOS 시소러스 • OSTI 시맨틱 시소러스 • EuroVoc
skos:prefLabel	• UNESCO 명명법 • UNESCO 시소러스 • Realfagstermer • PLOS 시소러스 • OSTI 시맨틱 시소러스 • EuroVoc

공통 핵심 요소	활용 체계
skos:altLabel	• UNESCO 시소러스 • Realfagstermer • PLOS 시소러스 • EuroVoc
skos:broader	• UNESCO 명명법 • UNESCO 시소러스 • Realfagstermer • PLOS 시소러스 • OSTI 시맨틱 시소러스 • EuroVoc
skos:narrower	• UNESCO 명명법 • UNESCO 시소러스 • PLOS 시소러스 • OSTI 시맨틱 시소러스 • EuroVoc
skos:related	• UNESCO 시소러스 • Realfagstermer • PLOS 시소러스 • OSTI 시맨틱 시소러스 • EuroVoc
skos:hasTopConcept	• UNESCO 명명법 • UNESCO 시소러스 • PLOS 시소러스 • OSTI 시맨틱 시소러스 • EuroVoc
skos:inScheme	• UNESCO 명명법 • UNESCO 시소러스 • OSTI 시맨틱 시소러스 • EuroVoc
skos:topConceptOf	• UNESCO 명명법 • UNESCO 시소러스 • OSTI 시맨틱 시소러스 • EuroVoc
skos:ConceptScheme	• UNESCO 시소러스 • Realfagstermer • PLOS 시소러스 • OSTI 시맨틱 시소러스
skos:scopeNote	• UNESCO 시소러스 • PLOS 시소러스 • OSTI 시맨틱 시소러스 • EuroVoc

〈표 11〉의 SKOS 요소들 중 Concept, prefLabel, altLabel은 용어를 표현하기 위한 가장 기본적인 요소로써, Concept는 용어가 표현하는 개념

을 나타내며, perfLabel은 이 개념을 대표하는 대표명을, altLabel은 대체명을 기술한다. 각각의 용어체계들에서는 보통 prefLabel과 altLabel에 영어, 스페인어 등 다양한 언어 표현을 언어 코드로 구분하여 함께 표현하고 있다. broader, narrower, related는 용어 간의 관계를 표현하기 위한 것으로 상위, 하위, 관련(특정 그룹에 함께 포함되거나 관련된 용어) 용어에 적용하며, 경우에 따라서는 분류체계에서의 상-하위 관계를 표현하는 데 이 세 요소를 사용하기도 한다. 이 외에 hasTopConcept, inScheme, topConceptOf, conceptScheme은 해당 용어가 속하는 범위를 지정하는 용도로 주로 사용되는데, 각각의 용어가 어떤 분류체계에 속하는지 표현하는데 있어 broader, narrow, related 대신 상위 분류체계 표현을 하는데 활용할 수 있다. 마지막으로 scopeNote는 주기사항, 즉 용어와 관련된 여러 가지 참고사항들을 기록하기 위한 요소이다.

4.2 표준분류체계 소분류 용어 관리를 위한 메타데이터 요소 검토

현재 표준분류체계의 경우, 소분류를 기술키워드로 대체·적용하기 위한 구체적인 실행 방안을 정립 중인 상황이므로 아직까지 용어관리에 대한 관리체계와 시스템이 마련되어 있지 않은 상황이다. 현재의 표준분류체계(과학기술정보통신부, 한국과학기술기획평가원, 2018b) 역시 책자 형태로 발간되어 있는 표준분류체계 연구분야 해설서(과학기술정보통신부, 한국과학기술기획평가원, 2018a)를 제외하고는 별도의 성

문화된 분류체계 관리 절차나 별도의 관리 시스템이 체계화되어 있지 못하다. 따라서, 표준분류체계 소분류 용어 관리를 위해 필요한 메타데이터 요소들을 도출하기 위해, 현재 제공되고 있는 표준분류체계 해설서의 내용을 검토해보고 공공데이터포털에서 서비스되고 있는 표준분류체계 데이터셋의 관리 요소들을 분석해 앞서 정리한 핵심 공통요소와의 비교를 통해 제안해 보고자 한다.

2018년 발간된 국가과학기술표준분류체계 연구분야 해설서(과학기술정보통신부, 한국과학기술기획평가원, 2018a)는 총 33개 대분류, 371개 중분류, 2,898개 소분류로 구성된 2018년 기준 표준분류체계에 대해 중분류 내용을 기준으로 각각의 분류에 대한 정의와 포함되는 소분류의 내용을 설명하는 방식으로 정리되어 있다(<그림 1> 참조).

대분류에 대해서는 별도 설명이 없으며, 각각의 분류는 분류 수준별로 영문과 숫자가 혼합된 코드값으로 식별할 수 있다. 분류 명칭은 한글명과 영문명이 제공되며, 중분류와 소분류에 대해서는 대략적인 정의 및 사용 범위에 대한 설명이 정리되어 있다. 이러한 해설서 구성체계의 내용들을 바탕으로 추출 가능한 메타데이터 요소들을 정리해 보면 <표 12>와 같이 나타낼 수 있다.

공공데이터포털¹⁾에서 확인할 수 있는 표준분류체계 데이터셋의 경우도, 해설서 내용을 통해 도출할 수 있는 요소와 크게 다르지는 않다. 한국과학기술기획평가원에서 제공하는 표준분류체계 관련 데이터셋은 CSV 파일 형태로 다운로드

1) <https://www.data.go.kr>



〈그림 1〉 표준분류체계 해설서 표지 및 대분류(수학), 중분류(응용수학) 해설 부분

〈표 12〉 해설서 기준 메타데이터 요소 도출

요소명(한글)	정의	비고
수준	분류 용어의 수준	대분류, 중분류, 소분류로 구분
코드값	용어의 식별자	영문과 숫자 조합으로 구성
한글명칭	용어의 한글명칭	
영문명칭	용어의 영문명칭	
정의	용어의 정의(대분류 제외)	대분류 용어의 경우 정의가 존재하지 않음
관계	용어 간의 상하관계	
상태	용어 상태	삭제 여부

드 받을 수 있으며, 2021년 11월 현재 ‘한국과학기술기획평가원(KISTEP)_과학기술표준분류_20210923.csv(이하, 과학기술표준분류 csv)’와 ‘한국과학기술기획평가원(KISTEP)_과학기술표준분류정보_20210923.csv(이하, 과학기술표준분류정보.csv)’라는 2가지 파일로 제공된다. 첫 번째 파일은 〈그림 2〉와 같이 ‘상위과학기술분류코드, 상위과학기술분류코드한글명, 과학기술분류코드, 분류명, 분류순서, 비고’의 칼럼명을 갖는 형태로, 두 번째 파일은 〈그림 3〉과 같이 ‘분류정보일련번호, 분류구분코드, 분류구분코드한

글명, 분류코드, 분류코드 한글명, 분야코드, 분야코드 한글명, 대분류코드, 대분류코드 한글명, 중분류코드, 중분류코드 한글명, 제목, 내용태그’의 칼럼을 갖는 형태로 구성된다.

공공데이터포털에서 다운로드 가능한 표준분류체계 관련 데이터셋의 칼럼명을 기준으로 메타데이터 요소들을 정리해 보면 〈표 13〉과 같다.

하지만 실제 입력된 데이터를 살펴보면 서로 칼럼명 내에 동일한 데이터 값들을 갖고 있는 것이 확인된다. 〈표 13〉에서 ‘과학기술표준분류

상위과학기술분류코드	상위과학기술분류코드한글명	과학기술분류코드	분류명	분류순서	비고
FSC		CL001	분류체계	1	분류정보, 임시분류 구분코드
CL001	분류체계	RCSARE	연구분야	1	국가과학기술표준분류체계_분류
RCSARE	연구분야	NAARE	자연	2	국가과학기술표준분류체계_분야
NAARE	자연	NA	수학(Mathematics)	1	국가과학기술표준분류체계_대분류
NA	수학(Mathematics)	NA01	대수학(Algebra)	1	국가과학기술표준분류체계_중분류
NA	수학(Mathematics)	NA02	해석학(Analysis)	2	국가과학기술표준분류체계_중분류
NA	수학(Mathematics)	NA03	위상수학(Topology)	3	국가과학기술표준분류체계_중분류
NA	수학(Mathematics)	NA04	기하학(Geometry)	4	국가과학기술표준분류체계_중분류
NA	수학(Mathematics)	NA05	응용수학(Applied Mathematics)	5	국가과학기술표준분류체계_중분류
NA	수학(Mathematics)	NA06	이산 및 정보수학(Discrete/Inform)	6	국가과학기술표준분류체계_중분류
NA	수학(Mathematics)	NA07	추론/계산(Parametric inference/c)	7	국가과학기술표준분류체계_중분류
NA	수학(Mathematics)	NA08	모델/자료분석(Models/Data anal)	8	국가과학기술표준분류체계_중분류
NA	수학(Mathematics)	NA09	응용통계(Applied Statistics)	9	국가과학기술표준분류체계_중분류
NA	수학(Mathematics)	NA10	확률/확률과정(Probability/Stocha	10	국가과학기술표준분류체계_중분류
NA	수학(Mathematics)	NA99	기타수학(other Mathematics)	11	국가과학기술표준분류체계_중분류
NAARE	자연	TA	수학	1	국가과학기술표준분류체계_대분류
TA	수학	TA01	대수학	1	국가과학기술표준분류체계_중분류
TA	수학	TA03	위상수학	2	국가과학기술표준분류체계_중분류
TA	수학	TA02	해석학	3	국가과학기술표준분류체계_중분류
TA	수학	TA04	기하학	4	국가과학기술표준분류체계_중분류
TA	수학	TA05	응용수학	5	국가과학기술표준분류체계_중분류
TA	수학	TA06	이산/정보수학	6	국가과학기술표준분류체계_중분류
TA	수학	TA07	추론/계산	7	국가과학기술표준분류체계_중분류
TA	수학	TA08	모델/자료분석	8	국가과학기술표준분류체계_중분류
TA	수학	TA09	응용통계	9	국가과학기술표준분류체계_중분류
TA	수학	TA10	확률/확률과정	10	국가과학기술표준분류체계_중분류
TA	수학	TA99	기타수학	11	국가과학기술표준분류체계_중분류

〈그림 2〉 ‘한국과학기술기획평가원(KISTEP)_과학기술표준분류_20210923.csv’ 파일 화면

분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역	분류코드	영역
------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----	------	----

〈그림 3〉 ‘한국과학기술기획평가원(KISTEP) 과학기술표준분류정보 20210923.csv’ 파일 화면

csv' 파일의 경우에는 해당 파일 안에서도 상위 과학기술분류코드와 과학기술분류코드, 그리고 상위과학기술분류코드명과 분류명에서 중복된 데이터가 존재하는데, 이는 해당 csv 파일 내용

이 대분류와 중분류에 해당하는 각각의 분류명을 기준으로 상-하위 개념의 연결 내용을 정리하였기 때문에 파악된다(〈그림 4〉 참조). ‘과학기술표준분류정보 csv’의 경우는 칼럼명이 대

〈표 13〉 공공데이터포털에서 제공되는 과학기술표준분류 데이터셋 파일 기준 관리요소 도출

구분	요소명(한글)	정의	비고
과학기술 표준분류 CSV	상위과학기술분류코드	상위분류체계 식별코드	영문2자리
	상위과학기술분류코드 한글명	상위분류체계의 한글 명칭	
	과학기술분류코드	용어의 식별자, 구분자	영문과 숫자 조합으로 구성
	분류명	용어의 한글명칭	
	분류순서	해당 용어의 분류체계 내 순서	
	비고	분류 용어의 수준	분야, 대분류, 중분류, 분류로 구분
과학기술 표준분류 정보 CSV	분류체계일련번호	일련번호	
	분류구분코드	분류구분 식별코드	영문2자리 숫자3자리 조합으로 구성되며 분 류정보와 임시정보를 구분함
	분류구분코드한글명	분류구분 한글 명칭	분류정보, 임시정보로 구분
	분류코드	분류 식별코드	영문6자리로 구성되며, 연구분야와 적용 분야 를 구분함
	분류코드 한글명	분류코드 한글 명칭	연구분야, 적용분야로 구분
	분야코드	분야 식별코드	영문과 숫자 조합으로 구성되며, 공공, 사회, 산업, 생명, 인간, 인간과학과 기술, 인공물, 자연의 정보를 구분함
	분야코드 한글명	분야코드 한글 명칭	공공, 사회, 산업, 생명, 인간, 인간과학과 기술, 인공물, 자연으로 구분
	대분류코드	대분류 식별코드	영문과 숫자 조합으로 구성
	대분류코드 한글명	대분류코드 한글명	표준분류체계 대분류명 기술(칼럼명은 한글 명이나 영문 명칭도 혼용되어 있음)
	중분류코드	중분류 식별코드	영문과 숫자 조합으로 구성
	중분류코드 한글명	중분류코드 한글명	표준분류체계 중분류명 기술(칼럼명은 한글 명이나 영문 명칭도 혼용되어 있음)
	제목	소분류 내용	칼럼은 제목으로 되어 있으나 표준분류체계 소분류에 해당하는 내용이 기술
	내용태그	별도 정의 없음	입력된 데이터 내용 없음

상위과학기술분류코드	상위과학기술분류코드한글명	과학기술분류코드	분류명
EA	기계(Machinery)	EA01	측정표준/시험평가기술(Meas
EA	기계(Machinery)	EA01	측정표준/시험평가기술(Meas
EA	기계(Machinery)	EA01	측정표준/시험평가기술(Meas
EA	기계(Machinery)	EA01	측정표준/시험평가기술(Meas
EA	기계(Machinery)	EA01	측정표준/시험평가기술(Meas
EA	기계(Machinery)	EA01	측정표준/시험평가기술(Meas
상위과학기술분류코드	상위과학기술분류코드한글명	과학기술분류코드	분류명
ATARE	인공물	EA	기계(Machinery)
ATARE	인공물	EA	기계(Machinery)
ATARE	인공물	EA	기계(Machinery)
ATARE	인공물	EA	기계(Machinery)
ATARE	인공물	EA	기계(Machinery)
ATARE	인공물	EA	기계(Machinery)

〈그림 4〉 과학기술표준분류 CSV 파일의 칼럼별 중복 데이터 입력값 사례

-중-소분류 체계로 구분되었기에 이에 따른 중복값은 보이지 않는다. 하지만 두 파일의 데이터들을 기준으로 보면 사실상 ‘과학기술표준분류정보 csv’의 분야코드, 분류코드, 대분류코드, 중분류코드와 각각의 한글명칭들은 ‘과학기술표준분류 csv’의 상위과학기술분류코드 및 과학기술분류코드와 각각의 명칭과 같은 내용을 담고 있으므로 사실상 하나의 메타데이터 요소로 통일해 정리가 가능하다.

이상에서 서술한 바와 같이 현재 표준분류체계와 관련해 제공되고 있는 해설서의 내용 및 데이터셋에서 제공되는 메타데이터 요소들 중 의미상 중복되는 요소를 통일한 후, 앞서 도출한 해외 주요 용어관리체계에서 활용하는 관련 요소들과 매핑해 정리해보면 <표 14>와 같이 정리해 볼 수 있다.

하지만 <표 14>에서도 표준분류체계 관리와 관련된 메타데이터 요소와 해외 사례에서 활용 중인 요소가 1:1로 매핑이 되지 않는다는 것은 해외 사례의 경우는 각각의 체계 안에서 구성되는 용어 하나하나를 의미 단위로 관리하는 반면,

표준분류체계에서는 분류체계라는 구조에 중심을 두고 데이터를 관리하기 때문이다. 따라서 앞으로 표준분류체계의 소분류를 기술키워드화해 용어 집합으로 관리하기 위해서는 분류체계의 한 구조로써가 아니라 각 용어 단위별로 개념을 표현하는 집합으로 관리가 필요하며, 그 내용이 현재의 중분류와 대분류 체계에도 연계될 필요가 있다. 때문에 본 연구에서는 표준분류체계 개정 작업의 단계별로 현재 체계에서 도출 가능한 메타데이터 요소들에서 작업을 시작해 점차 기술키워드의 개념 관리까지 확장하는 방식으로 메타데이터 요소를 확장·구축해 나가는 방안을 제안하고자 한다.

이를 위해 개정 작업의 첫 번째 단계(1단계)에서는 현재의 표준분류체계 관리 요소들을 중심으로 <표 15>와 같이 메타데이터 요소들을 정리해, 현재 표준분류체계의 용어들을 먼저 정리할 것을 제안한다. 사실 <표 15>에 정리된 요소들도 위에서 정리한 바와 같이 현재 표준분류체계에서 도출해 적용이 가능한 요소들이지 현재 해당 시스템에서 운영되고 있는 것은

<표 14> 표준분류체계 관리를 위한 관련 요소 매핑

요소명	정의	비고	매핑 가능 요소
코드값	용어의 식별자	영문과 숫자 조합으로 구성	dc:identifier
수준	분류 용어의 수준	대분류, 중분류, 소분류로 구분	skis:hasTopConcept, skos:inScheme, skos:topConceptOf, skos:conceptScheme
한글명칭	용어의 한글명칭		skos:Concept, skos:prefLabel skos:broader, skos:narrower, skos:related
영문명칭	용어의 영문명칭		skos:Concept, skos:prefLabel skos:broader, skos:narrower, skos:related
정의	용어의 정의	대분류 용어의 경우 정의가 존재하지 않음	skos:definition
관계	용어간의 상하관계		hasTopConcept, inScheme, topConceptOf, conceptScheme
상태	용어 상태	현재 사용하는지 여부 등을 표시	status, Previous_Classification, fstatus
비고	추가적인 정보	필요에 따라 내용 추가 가능	skos:scopeNote

〈표 15〉 표준분류체계 용어 관리 방안을 체계화 하기 위한 1단계 메타데이터 요소

요소명(한글)	정의	기술방법	제안 요소	비고
식별자	개념의 식별자	영문과 숫자 조합으로 구성	dc:identifier	
소속분류체계	해당 용어가 속하는 분류 체계 명칭	과학기술표준분류체계 명칭을 그대로 기술	skos:topConceptOf	한글과 영어 언어구분 코드로 모두 표기
용어수준	분류 용어의 수준	대분류, 중분류, 기술키워드 중 선택	termType	KISTEP 별도 정의
개념	분류체계 포함 용어가 상징하는 개념	용어의 개념	skos:Concept,	
용어명칭	용어명칭	특정 개념을 대표하는 명칭 한글을 기본으로 하되, 다양한 언어로 함께 병행하여 표기하고 국제표준 언어 구분기호를 활용하여 구분(ISO 639)	skos:prefLabel	1단계는 한글과 영어를 기본으로 함
용어정의	용어정의	분류체계 용어의 정의 한글을 기본으로 하되, 다양한 언어로 함께 병행하여 표기하고 국제표준 언어 구분기호를 활용하여 구분(ISO 639)	skos:definition	1단계는 한글과 영어를 기본으로 함
상위용어	특정 용어의 상위 개념 (상위분류)	특정용어(개념)의 상위개념을 연결하여 보여줌	skos:broader	기술키워드가 속한 중분류(복수선택 가능), 중분류가 속한 대분류 관계
하위용어	특정 용어의 하위 개념 (하위분류)	특정용어(개념)의 하위개념을 연결하여 보여줌	skos:narrower	대분류에 속하는 중분류, 중분류에 속하는 기술키워드(복수선택가능) 관계
상태	용어 상태	별도 입력값 정의 필요	status	KISTEP 별도 정의
최초생성일	해당 용어 최초 생성일	해당용어 최초 생성일	dc:created	
최근수정일	해당 용어 최근 수정일	해당용어 최근 수정일	dc:modified	

아니다. 따라서 1단계에서는 최대한 현재 도출이 가능한 요소들을 중심으로 관리 체계를 변화시키는 데 목적을 두었다.

〈표 15〉에서 ‘용어수준’의 경우 현재 표준분류체계 해설서나 데이터셋 모두에서 확인 가능한 요소이지만 별도로 관리되고 있다고 보기는 어렵다. 때문에 이를 대분류, 중분류, 기술키워드(소분류)로 명확하게 구분해 주는 것이 필요하다. 여기에 더 추가해야 할 요소는 ‘개념’이다. 사실 분류 대상 용어는 인간이 공유할 수 있는 특정한 개념과 의미를 언어의 형태로 표현한 것에 불과하다. 즉, 용어를 관리한다는 것은 특

정 언어로 표현된 하나의 기술용어를 관리하는 것이 아니며, 그 용어가 의미하고 있는 개념을 관리한다고 봐야 한다. 이 개념은 다양한 언어(예: 한국어, 영어, 일본어 등)와 표현 방식으로 관리하고 서비스할 수 있다. 현재의 표준분류체계에는 개념에 대한 데이터가 별도로 존재하지는 않으나 현재 관리하고 있는 언어로 정의할 수 있는 개념으로 접근이 가능하기 때문에, 개념을 관리함에 있어 한글이나 영어 표현 혹은 용어 식별자와 URI를 결합하는 등의 형태로 표현할 수 있다. ‘용어명칭’과 ‘용어정의’는 현재 내용 그대로 활용 가능한 요소이지만,

대분류의 경우 정의가 누락되어 있으므로 이 부분은 1단계 사업 진행 시 신규 구축이 필요할 것이다. 상위용어와 하위용어는 관계를 표현하기 위한 것으로 현재 분류체계의 관계를 활용해 두 용어 간의 관계 설정으로 쉽게 구현 가능하다. 단, 소분류 기술키워드로 사용되는 용어들의 경우에는 관계 설정을 중복 선택이 가능하도록 구성하여 향후 용·복합연구 현황 조사에 활용이 가능하도록 해야 할 것이다. ‘최초생성일’과 ‘최근수정일’ 역시 현재 존재하는 정보로 입력 가능한 요소이다.

‘상태’의 경우 현재 체계에서는 삭제된 용어에 대해서만 표현이 가능하도록 되어 있으나 장기적인 용어 관리 관점에서는 이보다 상세한 상태 정보 관리가 필요하다. 이러한 구체적인

용어 관리 상태를 표현하기 위해서는 2장에서 소개한 ISO/IEC 11179 메타데이터 레지스트리 표준 Part 6에서 제공하는 데이터 상태 정보(〈표 3〉 참조)를 바탕으로, ‘삭제’에 한정된 용어관리 상태를 좀 더 세밀하게 관리하기 위해 〈표 16〉과 같이 구체화할 것을 제안한다.

1단계의 메타데이터 요소 작업이 완료되면, 두 번째 단계(2단계)에서는 표준분류체계 용어들의 활용성 및 용어관리의 품질 자체를 높이기 위해 개별 용어를 중심으로 보다 많은 요소들의 입력을 추가로 시도할 필요가 있다. 이를 위해 SKOS 요소들 중 활용 가능한 요소들을 추가해 〈표 17〉과 같은 요소들을 확장 요소로 입력할 것을 제안한다.

〈표 16〉, 〈표 17〉에서 제안한 요소들은 실제

〈표 16〉 표준분류체계 ‘상태’ 정보를 구체화하기 위한 제안 내용

상태	코드값	설명
표준용어	Standard	표준분류체계 용어 심의를 마친 공식 용어로 현재 활용 가능한 용어
후보용어	Candidate	위원회의 심의 과정 중에 있는 용어로 향후 변경 가능한 잠재적 표준용어
비표준용어	Incomplete	향후 표준용어 활용 대상으로 신청된 용어
폐기용어	Retired	더 이상 사용하지 않는 용어

〈표 17〉 표준분류체계 용어 활용성 및 품질을 향상시키기 위한 2단계 확장 메타데이터 요소

요소명(한글)	정의	기술방법	제안 요소	비고
대체어	표준 선정 용어 외에 이를 표현하기 위한 대체 용어	약어, 철자변형 표현 용어 등을 기술	skos:altLabel	
이력주기	표준 선정 용어의 이력변경사항을 기록	개념의 과거 상태에 대한 용도 및 의미를 메모	skos:historyNote	
수정주기	표준 선정 용어의 수정사항에 대한 기록	개념 수정에 대한 참고 사항을 기술	skos:changeNote	
관련어	표준 선정 용어 중 상호간에 의미적 연관성이 높은 관련어를 연결지어 표현	관계표현으로 관련어와의 연결을 통해 기술	skos:related	
이전용어	현재는 표준 용어가 아니지만 과거에 사용했던 용어로 이동하기 위한 표현	관계표현으로 현재 비표준 용어를 새로운 용어로 이동시키기 위한 요소	previouslyTerm	KISTEP 별도 정의

※ skos요소 외 제안한 요소는 ‘이전용어’로 기존 용어가 폐기 혹은 더 이상 사용되지 않을 경우 이를 안내하기 위한 요소임.

용어관리시스템 구축에 활용할 수 있다. SKOS를 기반으로 하고 있다는 점에서 온톨로지 기반으로 데이터 모델링을 위해 필요한 클래스, 속성 명세는 <표 18>, <표 19>와 같이 정리된다. 단, <표 18>의 클래스 요소 중 skos:ConceptScheme은 제안요소가 아니지만, skos:topConceptOf 속성을 표현하기 위해서 활용해야 하는 표준용어로 포함시켰다. 같은 관점에서 <표 19>의 skos:hasTopConcept로 속성으로 포함시켰다.

이상 제안한 내용의 핵심은 현재 표준분류체계를 국제표준인 SKOS에 따라 용어관리체계로 전환하기 위해 우선 수행할 내용에 해당한다. 관리체계를 용어관리 관점으로 전환하면

해당 용어에 대한 관리체계가 수립되며, 이는 용어의 일부 혹은 전체로 데이터셋으로 활용할 수 있다. 또한 이를 기반으로 용어관리, 서비스 시스템을 구축함으로써 효율적인 용어관리가 가능해 용어의 품질과 활용성이 높아지는 것을 기대할 수 있다.

5. 결론 및 제언

현 표준분류체계를 분류구조를 단순화하고 소분류를 기술키워드로 대체하는 방향으로 개정하려는 주요 이유는, 분류체계가 국내외 R&D

<표 18> SKOS 기반 과학기술표준용어에 대한 클래스 명세

클래스명	정의	비고
skos:ConceptScheme	과학기술표준분류체계 명칭을 그대로 기술	
skos:Concept	용어의 개념	

<표 19> SKOS 기반 과학기술표준용어에 대한 속성 명세

속성명	유형	domain	range	비고
skos:topConceptOf	datatype	skos:Concept	skos:ConceptScheme	
skos:hasTopConcept	datatype	skos:ConceptScheme	skos:Concept	
kistep:termType	dataType	skos:Concept		list형으로 지정
dc:identifier	dataType	skos:Concept	xsd:string	
skos:prefLabel	dataType	skos:Concept	xsd:string	
skos:alt:Label	dataType	skos:Concept	xsd:string	
skos:definition	dataType	skos:Concept	xsd:string	
skos:broader	objectType	skos:Concept	skos:Concept	'skos:narrower'의 'inverseOf' 유형
skos:narrower	objectType	skos:Concept	skos:Concept	'skos:broader'의 'inverseOf' 유형
skos:related	objectType	skos:Concept	skos:Concept	
kistep:status	dataType	skos:Concept		list형으로 지정
dc:created	dataType	skos:Concept	xsd:date	
dc:modified	dataType	skos:Concept	xsd:date	
skos:historyNote	dataType	skos:Concept	xsd:string	
skos:changeNote	dataType	skos:Concept	xsd:string	
kistep:previouslyTerm	objectType	skos:Concept	skos:Concept	

관련 통계 분석부터 국가연구개발사업 운영 관리, 예산 배정 등 다양한 분야에서 중요한 기초 자료로 사용되고 있음에도 불구하고 각 분야 유관 분류체계와의 연계가 쉽지 않고 현재의 개정 주기로는 신기술 분야들을 유연하게 포용할 수 없는 경직된 구조를 갖고 있기 때문이다.

본 연구는 2022년 개정을 앞두고 있는 표준 분류체계의 개정 방향과 목적에 맞춰, 기존에 4단계 분류체계에서 최하위 계층으로 관리되고 있던 소분류를 기술키워드 방식으로 대체·확장해 활용하기 위한 개정 단계별 메타데이터 요소를 도출해 제안하고자 하는 목적으로 진행되었다. 이를 위해 SKOS와 ISO/IEC 11179 국제 표준의 내용을 검토하여 소분류 체계를 기술키워드 방식으로 대체 운영하게 되었을 때, 지속적으로 확장될 기술키워드 용어들을 효과적으로 관리하는 데 필요한 메타데이터 요소들을 확인하였다. 그 다음 과학기술 분야와 관련된 전문용어들을 관리하고 있는 해외 주요 용어관리 체계 사례들을 검토해, 현재 각 용어관리체계에서 활용 중인 메타데이터 요소들을 분석해 정리하였다. 그리고 마지막으로, 현재 표준분류체계에서 제공하고 있는 해설서 내용과 공공데이터 포털에서 서비스 중인 표준분류체계 데이터셋의 내용을 분석해 개정 작업 단계별로 소분류 기술키워드 용어 관리에 필요한 메타데이터 요소들을 도출해 제안하였다.

최종적으로 제안한 메타데이터 요소는 총 16가지로 표준분류체계 개정을 위한 2가지 작업 단계를 고려하였다. 1단계 작업 시 제안한 11가지 요소(〈표 15〉 참조)는 현재 표준분류체계의 내용을 통해 도출 가능한 메타데이터 요소들을 정리한 내용으로 현재의 소분류 용어를 기술키

워드로 변환하는 데 필요한 요소들이다. 그리고 2단계 작업 시 제안한 5가지 요소(〈표 17〉 참조)는 2022년 개정 이후 지속적으로 표준분류체계에 추가·갱신될 기술키워드 용어들의 개념 및 이력 관리에 대한 부분을 고려해 필요한 메타데이터 요소들을 제안하였다.

물론 본 연구에서 제안한 메타데이터 요소들은 실제 적용 가능성 여부를 판단하기 위한 심화 연구가 필요하며, 관점에 따라서는 용어 관리에 필요한 요소들이 더 요구될 수도 있다. 또한 소분류 용어들을 기술키워드 방식으로 변경 관리하기 위해 진행해야 할 현실적인 작업 과정 역시 이제부터 시작이다. 운영조직 개편부터 현 분류체계를 기술키워드 방식을 적용한 체계로 재구축하기 위한 관리시스템의 재설계, 향후 용어를 확장 구축하는데 필요한 관련 데이터 확보 등이 그것이다. 하지만 앞서도 언급했듯이 용어 관리에 활용할 수 있는 메타데이터 요소들은 각 시스템의 목적에 따라 각각 다르기 때문에, 단순히 관리하는 요소의 숫자가 많고 상세하다는 것만으로 용어 관리의 품질이나 우수성을 판단할 수는 없다. 그럼에도 불구하고 표준분류체계와 같이 국내외 주요 분야에서 다방면으로 활용되는 기초데이터 성격의 용어라면, 다양한 이용자들의 편의성과 요구사항들을 충족시킬 수 있는 메타데이터 요소들을 충분히 고려해 반영할 필요도 있다. 또한 33개 대분류, 371개 중분류, 2,898개 소분류 용어들을 관리하는 현재의 분류체계 방식을 국제 표준에 적합한 용어관리 방식으로 개편하는 과정에 있어서의 효율성 문제도 고려해야 한다.

따라서 본 연구는 현재 분류체계를 개편하는

데 있어 작업 과정의 혼란을 최소화하고 효율적인 개정 작업이 이루어질 수 있도록 현재의 표준분류체계에서 제공되는 자료들을 바탕으로 적용할 수 있는 메타데이터 요소들을 도출해 제안했다는 데 그 의미가 있다. 더불어 국내외 관련 표준과 사례들을 분석해 1단계 변환 작업 완료 후 지속적으로 갱신되는 용어들을 체계적으로 관리하기 위해 2단계 확장 메타데이터 요소를 추가 제안했다는 점도 의의가 있다.

하지만 기존에 표준분류체계가 활용되었던 전 분야에서, 개정 이후 소분류 기술키워드 용어가 무리 없이 적용될 수 있도록 하기 위한 후속 연구와 관련 작업은 앞으로 이어져야 할 필요가 있다.

참 고 문 헌

- 과학기술기본법. 법률 제18069호.
 과학기술기본법시행령. 대통령령 제32137호.
 과학기술정보통신부 (2019). 국가과학기술표준분류체계 개정 타당성 평가 대상 선정결과(안) 및 중·장기 개선방향.
 과학기술정보통신부, 한국과학기술기획평가원 (2018a). 2018년 국가과학기술표준분류체계 연구분야 해설서.
 과학기술정보통신부, 한국과학기술기획평가원 (2018b). 국가과학기술표준분류체계(2018년 개정).
 최문정, 임현, 이상남, 박노언, 정의진, 안지현, 조아라, 이해림, 지수영 (2019). 과학기술기획 및 혁신 정책 활용도 제고를 위한 KISTEP 미래예측 역할 재정립 연구 (기관-2019-048). 한국과학기술기획평가원.
 ISO/IEC JTC1 SC 32 WG2 (2015, November 5). ISO/IEC 11179, Information Technology - Metadata registries (MDR).
 Available: <https://metadata-standards.org/11179/index.html>
 Miles, A. & Brickley, D. (2005, November). SKOS Core Vocabulary Specification. W3C.
 Available: <https://www.w3.org/TR/swbp-skos-core-spec/>

• 국문 참고자료의 영어 표기

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

- Choi, M., Im, H., Lee, S., Park, N., Jeong, E., Ahn, J., Jo, A., Lee, H., & Ji, S. (2019). Study on Re-establishment for Future Prediction Roles of KISTEP to Improve Utilization of

Science and Technology Planning and Innovation Policy (Institution-2019-048). Korea Institute of S&T Evaluation and Planning.

Enforcement Decree Of The Framework Act On Science And Technology. Presidential Decree No. 32137.

Framework Act On Science And Technology. Act No. 18069.

Ministry of Science and ICT, Korea Institute fo S&T Evaluation and Planning (2018a). 2018 National Science and Technology Standard Classification System Research Fields Guideline.

Ministry of Science and ICT, Korea Institute of S&T Evaluation and Planning (2018b). National Science and Technology Standard Classification System(revised in 2018).

Ministry of Science and ICT, Korea Institute fo S&T Evaluation and Planning (2019). The selection results(draft) of targets for the revision feasibility evaluation about the National Science and Technology Standard Classification System, and the improvement directions in the mid to long term.