한의학 연구 데이터 관리 및 공유를 위한 메타데이터 요소 설계*

Metadata Element Design for Korean Medicine Research Data Management and Re-use

예 상 준 (Sang-Jun Yea)**
장 호 (Ho Jang)***
김 선 태 (Suntae Kim)****

목 차

1. 서 론
2. 이론적 배경
3. 메타데이터 요소 설계를 위한 사전 검토
4. 메타데이터 요소 설계 및 검증
5. 결론 및 제언

초 록

본 연구에서는 한의학 연구 데이터 관리 및 공유를 위한 메타데이터 요소를 설계하고 한국한의학연구원에서 생산되는 연구 데이터를 대상으로 요소검증을 수행하였다. TTAK.K0-10.0976 표준과 DataCite 메타데이터 스키마, 한국과학기술정보연구원의 국가 연구 데이터플랫폼 스키마를 대상으로 메타데이터 요소를 도출하였다. 최상위 요소로서 Identifier를 포함한 27개 요소가 도출되었으며, 필수요소 29개, 권고요소 13개, 선택요소 31개를 도출하였다. 한의학 관련 분야별 전문가 6인을 대상으로 도출요소에 대한 필요성 정도를 조사하였으며, 추가적인 메타데이터 요소를 제안 받아 검토하였다. 한의학 연구 데이터 주제 분류를 제안하였으며, 최종 제안하는 메타데이터 요소로 현재 운영 중인 시스템의 메타데이터 수용여부를 검증하였다. 본 연구결과는 향후 한국한의학연구원에 구축되는 데이터 리포지터리에 아이템과 파일을 기술하기 위한 메타데이터 스키마로 활용될 수 있다.

ABSTRACT

This research makes the metadata element design for Korean medicine research data management and re-use. Derived metadata elements are verified in research data of Korea Institute of Oriental Medicine. TTAK.K0-10.0976 Standard, DataCite metadata Schema and National Research Data Platform of KISTI were analyzed to derive the metadata elements. Including Identifier, 27 elements were derived as top-level elements with 29 mandatory elements, 13 recommended elements and 31 optional elements. The degree of elements' necessity and new metadata elements were investigated and suggested in the survey by six domain experts in korean medicine field. In this study subject classification for the korean medicine research data are suggested. The final version of metadata schema was tested and verified by comparing with the legacy metadata fields. The research results can be used to describe the Korean medicine research data: items and files.

키워드: 연구 데이터, 한의학, 메타데이터, 리포지터리, 한국한의학연구원
Research Data, Korean Medicine, Metadata, Repository, Korea Institute of Oriental Medicine

---

* 본 연구는 2019년도 전북대학교 연구기반 조성비 지원을 받아 수행되었음.
본 연구는 한국한의학연구원 '연구정보 빅데이터화 및 통합 관리체계 구축(KSN1812640)' 과제의 지원을 받아 수행되었음.
** 한국한의학연구원 선임연구원(tomita@kiom.re.kr / ISNI 0000 0004 6775 3197) (제1저자)
*** 한국한의학연구원 선임연구원(jh@kiom.re.kr / ISNI 0000 0004 7640 9793) (공동저자)
**** 전북대학교 문헌정보학과 조교수(kim.suntae@jbnu.ac.kr / ISNI 0000 0004 6492 6355) (교신저자)
논문접수일자: 2019년 4월 14일 최초심사일자: 2019년 5월 7일 게재확정일자: 2019년 5월 20일
한국문헌정보학회지, 53(2): 223-246, 2019. 〔http://dx.doi.org/10.4275/KSLIS.2019.53.2.223〕

www.kci.go.kr

1. 서 론

1.1 연구 필요성

연구 데이터 관리 및 공유와 관련된 활동이 활발하게 진행되고 있다. 국민 세금이 투입되어 생산된 연구 데이터를 국가 자산 및 기관 자산으로 보존하고 재사용 될 수 있도록 다양한 연구와 성공적 사례가 만들어지고 있다. 이러한 활동은 주요 선진국에서 오픈 데이터 정책을 이끌어내고 있으며, 연구 데이터 관리 및 서비스, 데이터 관리 계획(Data Management Plan, 이하 DMP)과 같은 구체적 사례를 만들어 내고 있다. 오픈 사이언스 실천운동으로 오픈 데이터 활동이 글로벌하게 확산되고 있는 것이다. 사회구성원 누구나 법적, 제도적, 기술적 장벽 없이 공공연구 성과물에 접근가능하고 이를 재사용 가능하도록 연구 환경이 진화하고 있는 것이다. 연구 데이터와 관련된 이러한 논의는 오래전부터 있어왔지만 연구 재현성과 복제성, 일반화 가능성에 대한 관심이 높아지면서 활발해졌다. 이와 관련하여 미국 국립 아카데미는 연구 데이터 완전성과 접근성, 관리성을 3대 관리원칙으로 제시하며, 이를 실천하기 위한 11개 권고안을 제안한바 있다(National Academy of Sciences, National Academy of Engineering and Institute of Medicine Committee 2009).

한편, 오픈 사이언스 시대에 맞추어 한의학 분야에서도 관련연구가 활발하다. 한의학 분야 연구 데이터는 고문헌의 분석, 천연물의 유효성 및 안전성에 대한 실험 데이터, 침술 및 부황 등과 같은 한의학적 치료법에 대한 연구 데이터, 한의학적인 치료 기전에 대한 연구 데이터 등 종류가 매우 다양하다. 그러나 이러한 연구 결과들은 최종결과물 형태인 논문 및 보고서 형태로 생성된 것들만 관리되어 활용되는 수준이다. 연구 과정 중에 산출되는 원 데이터 및 통계적 처리 등을 가하여 생성된 처리과정 데이터는 연구자가 개인적으로 보관하여 관리하므로 연구종료 후 손실 위험이 크다. 따라서 타 분야와 마찬가지로 한의학 분야에서도 연구 데이터가 체계적으로 보존되고 재사용 될 수 있도록 인트라가 필요하다. 한의학 분야 연구기관으로서 국내 최대기관인 한국한의학연구원에서는 한의학 분야의 연구 데이터를 체계적으로 관리하고 생산적으로 공유・활용하기 위한 인프라 구축을 계획하고 있다. 한의학 연구 데이터 인프라는 OAIS 참조표준모형을 준수해서 구축될 예정이다. 이에 따라 한의학 연구 데이터 수집과 보존, 검색과 출판을 위해, 각 부문별 메타데이터 요소 스키마 설계가 반드시 필요하다.

1.2 연구 목적 및 방법

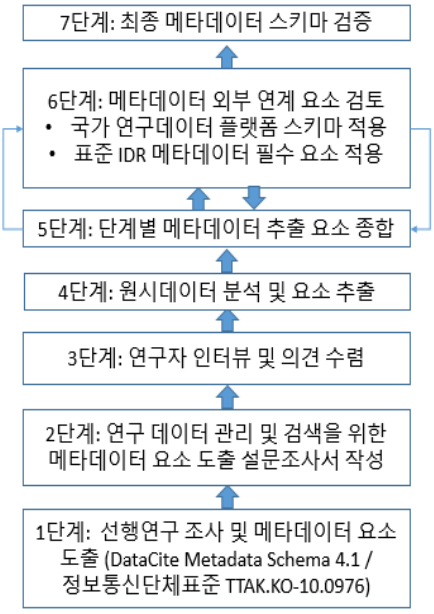
연구 데이터가 널리 활용되기 위해서는 데이터가 적절하게 기술되어 특정 연구 데이터를 원하는 이용자에게 쉽게 노출되도록 하는 작업이 필수적으로 요구된다. 즉 연구 데이터가 잘 관리되고 활용될 수 있는 적절한 메타데이터 요소를 선정하는 작업이 필요하다. 그러나 연구 데이터는 그 형식과 종류를 한정 지을 수 없으므로 논문과 같은 저작물들을 관리하기 위한 기존에 알려진 일부 메타데이터 요소들만으로는 충분히 기술될 수 없다. 예를 들어 유전체 염기서열 데이터라고 하면 염기서열 전 단계의

www.kci.go.kr

실험 조건이나 염기서열 데이터를 만들어주는 시퀀서 제품명 등도 메타데이터의 요소로써 활용이 될 수 있을 것이다. 더 나아가 한의학 도메인의 연구 데이터들은 그 도메인의 특성상 한약제제의 유효성 및 안전성 연구 고문헌의 분석 및 한의학적 치료법과 같은 특정한 주제에 특화되어 있으므로 그러한 도메인의 특성이 잘 반영되도록 메타데이터를 항목을 설계하는 것이 요구된다.

본 연구에서는 한국한의학연구원에서 구축 예정인 연구 데이터 리포지터리의 데이터 관리 구조를 설계하는데 있다. 특히, 한의학 연구 데이터 관리와 검색을 통한 재사용성을 높이기 위해서 관련된 메타데이터를 도출하고 도출된 요소를 통해 리포지터리에서 사용될 스키마를 제안하는데 그 목적을 둔다. 연구 범위로는 정보자원의 수집, 저장, 관리, 보존, 배포(출판)와 관련된 메타데이터만을 대상으로 요소 도출 및 스키마 설계를 진행하였다.

메타데이터 요소 도출을 위해 관련 표준들로 DataCite 스키마와 정보통신단체표준을 조사하였다. 이 과정을 통해 메타데이터 요소를 도출하였으며, 해당 내용을 바탕으로 연구자 인터뷰를 위한 설문지를 작성하였다. 연구자 인터뷰는 한의학 연구원의 핵심 연구자 6인을 대상으로 분야별 조사를 실시하였다. 인터뷰에 앞서 연구 데이터란 무엇인지? 설문조사는 왜 이루어지는지? 향후 설문결과가 어떻게 사용되는지에 대해, 인터뷰 전 1시간동안 인터뷰 대상자 교육을 실시하였다. 이는 인터뷰 과정에서 연구자의 적극적인 참여를 유도하기 위함이었다. 이러한 과정을 거쳐 수렴된 연구자 의견을 메타데이터 요소에 반영하였다. 〈그림 1〉은 본 연구에서 수행된 메타데이터 스키마 생성 절차를 보여준다. 4단계에서는 한의학연구원에서 수집, 생산되는 원 데이터를 분석하여 요소를 추출하였다. 5단계에서는 이전 단계에서 추출된 메타데이터 요소들을 종합하였다. 6단계에서는 메타데이터 외부 연계를 위한 요소를 검토하였다. 한국과학기술정보연구원에서 개발 중인 국가 연구 데이터 플랫폼 스키마와 표준 IDR(Institutional Data Repository)에서 요구하는 필수 및 권고 요소를 검토하였다. 마지막 7단계에서는 최종 메타데이터 스키마를 작성하여 검증하였다. 검증방법으로는 한의학연구원의 기존 시스템에서 연구 데이터 관리를 위해 사용 중인 메타데이터 항목을 개발된 메타데이터 요소와 매핑하여 누락되는 요소 없이 기술 가능한지 검토하는 방법을 적용하였다.



〈그림 1〉 메타데이터 스키마 생성 절차

www.kci.go.kr

2. 이론적 배경

2.1 연구 데이터

연구 데이터는 연구과정에서 생산되는 데이터로서 연구 결과를 재현, 복제, 일반화하는데 사용될 수 있는 데이터를 말한다. 좀 더 구체적으로는 논문이나 보고서와 같은 연구 출판물 내용을 증명할 수 있는 데이터로서 관찰, 관측, 조사, 실험, 분석 등과 같은 연구 활동을 통해 생산되는 데이터라 할 수 있다. 최근 NMC(The New Media Consortium)는 향후 5년간(2017년~2021년) 주요 흐름이 될 수 있는 6개의 핵심 트렌드와 6개의 주요 도전, 6개의 기술 발전에 대해 다룬 보고서, 'NMC Horizon Report: 2017 Library Edition'을 발표하였다. 보고서에서는 향후 1-2년 내 '연구 데이터 관리'를 위한 새로운 기술 도입이 필요하다고 분석하였다. 또한 향후 3-5년 이내 도서관 공간의 재구성을 위한 기술 도입이 필요하며, 장기적으로 학술 기록의 진화에 대비한 신기술 도입이 필요하다고 분석하였다(Becker et al. 2017). 보고서에 따르면 연구 데이터는 2014년부터 핵심 트랜드로서 지속적인 관심을 받고 있다.

한편 북미연구도서관협회(Association of Research Libraries, ARL)는 '연구 데이터 관리 서비스'를 주제로 보고서를 발행했다. ARL 회원 도서관을 대상으로 각 도서관의 연구 데이터 관리와 관련된 활동, 액세스 및 아카이빙 등에 대해 조사한 결과를 담고 있다. 또한 연구 데이터 정책, 데이터 보유 정책, 데이터 관리 계획 도구, 작업 기술서, 데이터 보존 웹 페이지, 인력자원 등 각 도서관에서 수행하고 있는 서비스 체제 및 재원 등에 대한 내용을 포함하고 있다(SPEC Kit 334).

국내에서는 2019년 3월 19일에 '국가연구 개발사업의 관리 등에 관한 규정'(공동관리규정, 대통령령)이 개정되어다. 이에따라 국가 R&D 연구과제에 데이터 관리 계획서 제출이 의무화되었다. 제출된 DMP는 과제 선정 시나 최종평가 시에 이행여부점검에 사용될 예정이다. 연구재단 및 정보통신기획평가원 연구과제에 시범적용하고 있으며, 국가 출연연구원에도 2020년 과제기획에 관련된 내용이 적용될 예정이다. 따라서 국내에서는 향후 5년은 연구 데이터 관련 이슈가 연구 커뮤니티에서 지속적으로 논의될 것으로 판단된다.

2.2 한의학 연구 데이터

다른 학문 분야와 유사하게 한의학에서도 연구 데이터에 대한 연구자간의 공감대 형성이 부족한 상황이어서, 한의연구원의 3개 연구부서장(미래의학부, 임상의학부, 한약연구부)의 추천을 통해, 분야별 전문가 20명을 선정하였다. 선정된 전문가 중에서 18명이 연구 데이터 자문회의에 참석하였다. 회의에서 전문가들의 논의를 거쳐서 확정된 한의학 분야의 연구 데이터에 대한 정의는 다음과 같다. "연구과정에서 생성된 실험・관찰・가공 데이터로써 원데이터 및 중간처리 데이터를 모두 포함하며 논문 및 보고서는 제외한다." 자문회의에서 전문가들이 제시한 한의학 주요 연구 분야에서 생성되는 데이터 유형은 〈표 1〉과 같다. 한의학 연구의 특성상 기초 생물 응용 연구에서부터 임상 연구를 거쳐 기기・정보 연구에 이르는

www.kci.go.kr

〈표 1〉 한의학 연구 분야별 데이터 유형

연구분야	데이터 유형	연구분야	데이터 유형
임상연구	한의코호트	한약재 효능 평가 (in vivo)	질환별(유도물질, 동물, 암수, 주령)
	임상시험 증례기록서		질환동물모델
의공학	뇌졸중 환자의 치료경과에 따른 요골동맥 맥파 데이터		질환별 동물모델
	갱년기증후군 치료전후에 따른 요골동맥 맥파 데이터	오믹스데이터	MS Data
	난임 환자 대상의 요골동맥 맥파 데이터		NMR Data
	대상증후군 태음인 대상의 요골동맥 맥파 데이터		임상조건
	뇌졸중 환자의 치료경과에 따른 생체 임피던스 데이터		gene sequence data
	당뇨환자의 생체임피던스 데이터	뇌영상데이터	뇌영상 자료
	상기도 감염 환자의 설 데이터	한약자원	증거표본 정보
생물학	유전체 칩		약재표본 정보
문헌정보	고문헌 관련 Text data		파생자원 정보
정보전산	시스템 기능 정의서		형태학적 감별정보
	프로세스 설계서		유전자 감별 정보
	UI 설계서		이화학적 감별 정보
	프로그램 소스		증식 기술 정보
	데이터베이스	한약 안전성 평가	in vivo 독성
한약재 효능 평가 (in vitro)	효소활성평가 데이터		in vitro 독성
	질환별 바이오마커		약물상호작용
	세포독성 검사		임상 연구
	단백질 발현 분석	한약 유효성 평가	한약 소재 효능 유효성 평가
	유전자 발현 분석	한약소재분석	한약소재 분석(1차생산)
	질환별 효능 평가 시스템		한약소재 분석(2차생산)

광범위한 분야에서 연구 데이터가 생성되는 것으로 분석되었으며 최근에 각광 받고 있는 오믹스 연구 및 뇌 연구에서 생성되는 데이터도 존재하는 것으로 밝혀졌다.

한의학 연구원에서 생산되는 연구 데이터는 크게 진단 연구 데이터, 약물치료 연구 데이터, 비약물치료 연구 데이터, 예방/협진 연구 데이터, 정보 연구 데이터, 정책/교육 연구 데이터로 구분될 수 있는 것으로 자문회의에서 협의되었으며 각각의 특징에 대해서 논의되었다. 〈표 2〉는 한의학 연구 데이터 구분과 각각의 특징을 보여준다.

〈표 1〉은 한의학 분야 연구 데이터 유형을 보여주며 〈표 2〉는 해당 분야 연구 데이터 구분과 각 각의 특징을 설명하고 있다. 한의학 분야 연구 데이터는 한의학・생물학・의학・정보학 등 다양한 학문분야와 연관되어 있으며, 데이터 유형 또한 기기생성 데이터・임상연구 데이터・유전자 데이터・정보전산 데이터 등 범위가 너무 커서 효율적인 관리가 매우 어려운

www.kci.go.kr

〈표 2〉 한국한의학연구원 연구 데이터 구분 및 특징

연구 데이터 구분	특징
진단 연구 데이터	사진(四診) 데이터, 변증진단 데이터, 체질진단 데이터, 진단기기 데이터 등을 의미
약물치료 연구 데이터	한약자원 데이터, 한약방제 데이터, 한약제품 데이터 등을 의미
비약물치료 연구 데이터	침/뜸/부항 데이터, 추나/도인 데이터, 경략경혈 데이터, 치료기기 데이터 등을 의미
예방/협진 연구 데이터	예방/예측 데이터, 채활치료 데이터, 양한방협진 데이터 등을 의미
정보 연구 데이터	문헌정보 데이터, 생물정보 데이터, 의료정보 데이터 등을 의미
정책/교육 연구 데이터	정책 데이터, 통계 데이터, 표준 데이터, 교육 데이터 등을 의미

상황이다. 이에 한의학 분야 연구 데이터가 체계적으로 관리되고 생산적으로 공유・활용되기 위해서는 체계적으로 구성된 메타데이터 스키마 설계가 필수적이다.

2.3 선행연구 분석

데이터가 없으면 과학이 진행될 것인가? 경험이 연구의 중심도구로 사용되던 1세대 연구에서부터, 이론이 중심도구로 사용되던 2세대 연구, 컴퓨터와 같은 계산 자원을 연구의 중심도구로 사용하는 3세대 연구, 데이터가 연구의 중심도구로 사용되는 4세대 연구 패러다임이 도래하기까지 연구자들은 끊임없이 데이터를 생산, 수집, 분석하며 연구를 수행하고 있다. 데이터 없는 연구는 불가능하다. 하드웨어와 소프트웨어, 네트워크 기술, 센서 기술 등이 발전하면서 데이터 중심의 과학시대가 도래하였다. 연구자들은 연구 결과의 진실성을 확보하고 연구 내용을 뒷받침할 수 있는 데이터와 데이터 재사용에 관심을 갖기 시작했다. 국가나 연구기금기관, 연구기관도 국가 자산, 기관자산으로서의 연구 데이터에 관심을 갖기 시작했으며, 체계적인 자산관리와 재사용성을 높이기 위한 노력을 기울이고 있다.

유럽의 경우, 'Open Innovation, Open Science, Open to the World' 비전을 선포하고 연구 데이터 호스팅과 접근, 거버넌스 체계를 지원하는 유럽 오픈 사이언스 클라우드 인프라스트럭처를 구축 중에 있다. 정책적 접근으로는 연구자 보상, 평가(almetrics), 유럽 오픈 사이언스 클라우드, 출판 비즈니스 모델 변경, 연구 진실성, 시민 과학, 열린 교육과 기술, FAIR 오픈 데이터와 관련된 이슈들을 논의하기 위한 오픈 사이언스 정책 플랫폼을 운영 중이다(European Commission 2016). 한편, 유럽 시장경제 활성화를 위해 추진하고 있는 Europe 2020 프로젝트와 세계적 수준의 과학실현과 시장경제 장애 요소를 극복하기 위해 추진하고 있는 Innovation Union 프로젝트를 성공적으로 수행하기 위해 추진하고 있는 연구기금 지원 프로그램으로 Horizon 2020 프로그램을 운영 중이다. 그 예산이 2014년부터 2020년까지 100조원에 달한다. 미국의 경우, 대통령 메모를 통해 연방정부 산하 기관과제에서 나온 출판물 및 데이터에 대한 오픈액세스를 권고하고 있다. 또한 디지털 과학데이터의 공개적 접근, 검색, 추출, 분석을 위한 데이터 관리 계획을 요청하고 있다(신은정 2018). 영국은 2016년 7월, HEFCE와 Rsearch Councils UK, Universities UK, Wellcome Trust

www.kci.go.kr

와 같은 다양한 연구 커뮤니티 이해관계그룹이 모여 연구 데이터 공개에 대한 협약내용을 발표하였다. 이 협약내용에서 연구 데이터란 인쇄든 디지털이든 물리적 형태를 취하든 연구 질문에 대한 해답을 뒷받침할 수 있는 증거로서의 데이터로서 발견을 검증하는데 사용될 수 있는 데이터(Rylance et al. 2016)로 정의하고 있다. 일본 과학기술국(Japan Science and Technology Agency, 이하 JST)은 ICT발달로 오픈 사이언스가 가속화 되고 있음을 밝히며 2017년 4월, 연구 출판물에 대한 오픈 액세스와 연구 데이터 관리에 대한 정책(JST 2017a)과 가이드라인을 발표하였다. 특히, 연구 데이터 관리 부문에 대해서는 연구 출판물의 기저가 되는 데이터 공개를 권고하였다. 데이터보존장소로 공공 데이터베이스와 분야별 리포지터리 등을 제안하고 있다. 또한 일본의 정보학연구원, NII (National Institute of Informatics, 이하 NII)에서 제공하는 리포지터리 리스트를 제시하고 있다. 또한 데이터 관리 계획(Data Management Plan, 이하 DMP)에서 일반적으로 요구되는 기술사항으로써 데이터 보존 및 관리에 관한 기본정보부터 데이터 제공방식 등의 가이드를 제안하고 있다(JST 2017b). Rolando et al. (2015)는 연구자들이 작성한 DMP를 분석함으로써, 연구자들이 무엇을 요구하는지? 데이터에 대한 연구자들의 인식 수준은 어느 정도인지? 도서관이 새로운 연구 데이터 관리 서비스(Research Data Management, 이하 RDM) 서비스를 어떻게 기획해야 하는지 아이디어를 얻을 수 있으며, DMP가 RDM 서비스 기획의 시작도구로 사용될 수 있다고 주장했다.

한편, 연구 데이터와 관련된 다양한 설문조사 연구도 수행되었다. Johnson et al. (2016) 연구에 따르면, 영국의 6개 고등교육 기관 연구자 약 1천명을 대상으로 한, 설문조사 결과를 통해 다음 사항을 알 수 있다. 디지털 연구 데이터 유형을 조사한 결과 문서 및 PDF 등의 데이터가 75% 이상을 차지하며, 다양한 연구 데이터 생산수단 및 데이터 형태가 존재한다. 또한 연구자들은 데이터 보존 장소로서 기관형 데이터 리포지터리(Institutional Data Repository, 이하 IDR이라함)와 연구 데이터 장기 보존을 요구하고 있다. 설문에 응답한 연구자 1,185명 연구자 중에 68% 응답자는 이미 연구 데이터를 공유하고 있거나 미래에 그러기를 희망한다고 응답하였다. 이러한 응답 요인으로는 '연구 데이터를 공적 자산으로 인식'과 데이터 재사용, 연구 결과 검증, 데이터 인용, 연구비 지원 기관 요구, 기관 데이터 정책 순으로 응답하였다. 한편 응답자 중 64% 연구자들은 기관의 연구 데이터 서비스를 사용하고 있지 않았으며, RDM 서비스 존재 자체를 모르는 연구자도 35%에 달했다. Tenopir et al. (2017)는 유럽 연구도서관협회(LIBER)에서 학술연구도서관들을 대상으로 연구 데이터 서비스에 관한 설문 조사를 진행하였다. 22개국 119개 기관이 응답하였다. 설문 결과를 분석한 결과, 도서관은 현재 기술형 서비스(데이터 저장 등) 보다 상담형 서비스(어떻게 정보를 찾는가, 데이터 관리 계획, 메타데이터 표준, 데이터 인용 등)를 제공하고 있다. 연구 데이터 서비스 정책을 가지고 있는 기관은 아직 절반 이하에 불과하며, 응답한 도서관장이나 정보센터 책임자 중 2/3가 도서관의 연구 데이터 서비스 필요성에 강하게 동의하고 있다고 조사되었다.

www.kci.go.kr

한편, 한국한의학연구원에서는 연구원 내부에서 생성되는 연구 데이터를 수집・관리하기 위한 선행 작업으로 “한의학 자원 DB 통합을 위한 정보화 전략 기획(ISP)에 관한 연구”를 수행하였다. 조성복 등(2012)은 개별 연구자가 보유하고 있는 Non-digital/Digital 연구 데이터를 수집하고 수집된 연구 데이터에 대한 범주화를 거쳐 한국한의학연구원에서 생성되는 연구 데이터 활용에 대한 요구사항을 분석하였다. 연구를 통해서 효과적인 연구 데이터 관리를 위한 유연한 조직구성과 조직문화가 필요함을 역설하였으며 적절한 보상체계가 마련되어야 함을 밝혔다. 또한, 연구 데이터 수집・관리 및 통합 DB 시스템을 관리 운영할 수 있는 조직과 지속적인 예산 지원이 필요함을 주장하였다.

3. 메타데이터 요소 설계를 위한 사전 검토

본 연구에서는 메타데이터 요소 값으로 사용될 통제어에 대한 조사를 실시하였다. 또한 연구 데이터 공유 및 관리를 위한 관련 표준으로써 정보통신단체표준 TTAK.K0-10.0976 및 DataCite 스키마를 검토하였다. 연구 데이터 관리 및 재사용성을 높이고 연구 데이터 관리 시스템 간 상호 운용성을 확보하기 위해 국내외 표준에 대한 사전검토를 수행하였다.

3.1 통제어 검토

한국한의학연구원에서 생산되는 연구 데이터 수집・공유를 위한 제도를 구축하고 관리・활용을 위한 리포지토리 시스템을 개발하기 위해서 한국한의학연구원에서 수행되는 주요사업 및 수탁과제에서 생산되는 원 데이터 및 중간처리 데이터를 분석하여 연구 데이터 분류체계를 개발하였다. 분류체계 초안 작성을 위해, 기존 연구결과(조성복 2012)와 현존하는 기술분류체계(국가과학기술표준분류(2018), 보건산업기술분류체계(2013), HT표준기술분류체계(2014))를 참고하였다. 또한 다양한 선행연구(Süt 2014; Zeng et al. 2015; Röhrig et al. 2011)를 참고하였다. 초안은 3회에 걸친 한의학 전문가 자문을 통해 수정 보완 되었다. 2.2절에서 밝혔듯이, 한의학 분야 연구 데이터는 유형과 형식이 매우 다양하다. 따라서 하나의 분류체계로 모든 데이터를 기술하는 것은 분류체계의 간결성 및 직관성을 훼손하여 분류체계 활용성을 저하시킨다는 전문가 의견을 수용하였다. 이에 분류체계는 연구주제, 연구방법, 질환/병증에 따른 세 가지의 분류체계로 구성되었으며, 하나의 연구 데이터에 세 가지 분류체계를 적용하여 자세한 연구 데이터 분류 태그가 생성될 수 있도록 고안하였다. 분류체계 특성은 〈표 3〉과 같으며, [부록 1]을 통해 세부 내용을 확인할 수 있다.

〈표 3〉 분류체계별 특성

분류체계	분류 깊이	분류 개수
연구주제	대-중-소	소분류 기준 108개
연구방법	대-중	중분류 기준 28개
질환/병증	대-중	중분류 기준 58개

www.kci.go.kr

본 연구를 통해 도출한 분류체계는 한의학 분야 연구 데이터 리포지터리가 구축될 경우, 컬렉션 구성에 활용될 수 있을 것이다. 또한 연구 데이터를 기술한 아이템의 주제 정보를 통일된 방식으로 기술할 수 있는 기준이 될 수 있다.

3.2 관련 표준 검토

3.2.1 정보통신단체표준 TTAK.K0-10.0976

TTAK.K0-10.0976은 메타데이터 프로젝트 그룹(PG606)에 제안되고, 소프트웨어/콘텐츠 기술위원회(TC6)에서 심의를 통과하여 2017년 3월에 승인된 정보통신단체표준이다. 표준 명칭은 '연구 데이터 관리 및 공유를 위한 메타데이터(The metadata for the managing and sharing research data)'이다.

연구 과제를 수행하며 생산되는 연구 데이터들을 메타데이터 기반으로 효율적으로 관리, 공유 및 활용 할 수 있도록 메타데이터 관리체계, 요소 및 세부 사항을 표준 (안)으로 명시하고 있다. 해당 표준은 연구 데이터의 메타데이터 표준은 연구과제에서 발생하는 연구 데이터를 효과적으로 관리, 공유하고 재활용함으로써 연구자 및 기관의 자산인 데이터를 보존하고 데이터의 재활용을 통해 연구의 효율성을 제고하는 도구로 활용될 목적으로 제안되었다(김선태 외 2017). 표준에서는 연구 데이터 관리를 위한 메타데이터 요소를 네 개 단위로 다음과 같이 명시하고 있다. 연구 데이터 관리 단위, 컬렉션 메타데이터 요소, 데이터셋 메타데이터 요소, 마지막으로 파일 메타데이터 요소로 구분되어 있다. 또한 리포지터리 관리를 위한 메타데이터 요소도 명시되어 있다.

3.2.2 DataCite 메타데이터 스키마 4.1

DataCite는 국제적인 컨소시엄으로서 2018년 12월 현재, 전 세계 1312개 데이터센터가 DataCite에 메타데이터를 출판하고 있다. DataCite 메타데이터 최신 커널의 주요한 기능은 크게 4가지로 요약할 수 있다. 첫째, DOI 등록을 위해 필요한 최소한의 요소를 이용해 데이터세트를 인용하기 위한 표준을 제시하는 기능, 둘째, 데이터 관리를 위한 다양한 스키마들과 상호운용성 확보 기능, 셋째, 다른 자원과의 관계를 기술하는 것과 같이 데이터에 대한 융통성 있는 기술이 가능하도록 부수적인 요소를 제공함으로써 데이터세트의 발견이 더욱 용이하도록 하는 기능, 넷째, 용어의 통제어 리스트 사용이나 표준 어휘집 사용을 통해 미래 서비스를 준비하는 기능으로 구분 된다(Starr and Gastl 2011).

DataCite 메타데이터 스키마 3.1 버전은 2회에 걸쳐 갱신이 이루어져 2017년 10월에 메타데이터 스키마 4.1 버전을 발표하였다. 이전 버전과 달라진 주요 변경사항을 살펴보면, 1) 'GeoLocation' 요소에 복수의 polygons 정보를 기술할 수 있도록 허용하고 있다. 2) polygon 관련 요소를 위한 하위 속성으로, 선택적 사용 가능한 'inPolygonPoint' 요소를 추가 하였다. 3) 통제어로 제공되는 'dateType'에 'Other'를 추가하였다. 4) 자원유형을 기술하기 위한 통제어휘 'resourceType'으로 'DataPaper'를 추가하였다. 5) 자원 사이의 관계를 기술하는 통제어로 'IsDescribedBy'와 'Describes', 'HasVersion'과 'IsVersionOf', 'IsRequiredBy'와 'Requires'를 추가하였다. 6) 'Date' 요소의 하위 속성으로 'dateInformation'을 추가하였다. 7) 'creatorName'과 'ContributorName' 요소의 선택적 속성으로

www.kci.go.kr

'nameType'을 추가하였다. 'nameType'으로 기술될 수 있는 값은 personal과 organizational로 통제된다. 8) 관련된 자원의 식별자를 기술하는 'relatedIdentifier'의 선택적 하위 속성으로 'resourceTypeGeneral'을 추가하였다. 기술될 수 있는 값은 기존에 통제어로 제공되는 'resourceTypeGeneral' 값과 동일하다. 9) 저작권을 기술하는 'Rights' 요소에 언어 정보를 선택적으로 기술하기 위한 'lang' 속성이 추가되었다.

4.0 버전의 주요 변경사항을 살펴보면, 1) 선택적 요소였던 'resourceTypeGeneral' 요소가 필수요소로 변경되었다. 이는 데이터 관리와 검색의 효율을 높이고자 함을 의미한다 할 수 있다. 2) 연구비 지원기관 정보를 기술할 수 있도록 'FundingReference' 요소가 추가되었다. 하위속성으로 'funderName', 'funderIdentifier', 'awardNumber', 'awardURI' 그리고 'awardTitle' 속성을 갖는다. 이와 관련하여 'contributorType' 속성으로 기술될 수 있는 통제어 값으로 'funder'가 삭제되었다. 3) 'creatorName'과 'contributorName'의 선택적 하위 속성으로 'familyName'과 'givenName'이 추가되었다. 4) 통제어로 정의된 'relatedIdentifierType'의 값으로 'IGSN'이 추가되었다. 5) 위치정보를 기술하는 'GeoLocation' 요소의 하위 속성으로 'geoLocationPolygon'을 추가하였으며, 'geoLocationPoint'와 'geoLocationBox' 요소 설명을 재정의 하였다.

본 연구에서 검토된 정보통신단체표준과 DataCite 메타데이터 스키마를 이용해 1차적인 메타데이터 요소가 도출되었다. 이후, 연구자 인터뷰를 통해 메타데이터 요소 추가 및 변경 작업이 이루어졌다.

4. 메타데이터 요소 설계 및 검증

4.1 메타데이터 요소 도출

정보통신단체표준 TTAK.K0-10.0976과 DataCite 메타데이터 스키마를 분석하여 식별자(Identifier) 요소를 포함하여 총 21개 요소를 도출하였다. 이후, 도출된 요소를 바탕으로 메타데이터 요소 도출 설문 조사서를 작성하였다. 설문 조사서 작성문항에 응답자가 수행하는 연구방법과 질환병증을 표기할 수 있도록 분류표를 작성하여 제시하였다. 연구방법을 표기하도록 한 이유는 특정 연구방법에서 필요로 하는 메타데이터 요소를 도출하고자 함이다. 질환병증을 표기하도록 한 이유도 이와 같다. 연구방법과 질환병증과 관련된 메타데이터 요소는 연구 데이터 제출 시점과 검색시점에 입력과 검색 포인트로 활용될 수 있기에 매우 중요하다. 메타데이터 요소 도출을 위한 연구자 인터뷰는 분야별(정보, 기기, 한약효능, 셀, 동물실험, 임상, 유전체, 문헌) 대표 연구자 6인을 대상으로 진행하였다. 보다 깊이 있는 인터뷰와 설문조사를 진행하기 위하여 인터뷰 대상자 교육을 위한 발표 후 인터뷰를 진행하였다.

설문 내용은 기 도출된 요소들에 대한 필수, 권고, 선택 여부를 5점 척도로 의견을 묻는 것과 연구 데이터 관계 기술을 위한 요소 제안, 추가적인 메타데이터 요소 제안으로 구성하였다. 사전 설계 된 메타데이터 요소에 대한 필수, 권고, 선택 여부와 추가적인 메타데이터 요소 제안은 한의학 연구 분야 특성을 반영하기 위함이다. 설문조사 결과, 1) 데이터 출판에 대한 이해 부족으로 연구 데이터 출판을 위한 필수

www.kci.go.kr

요소(생산자/creator, 출판자/publisher, 출판년도/pulication year)에 대해서 선택 및 불필요 요소로 응답한 경우가 존재, 2) 데이터 생산자(creator)를 선택 요소로 응답한 경우가 존재. 이는 데이터 생산자가 연구 데이터 검색 및 재사용에 있어 반드시 필요한 요소가 아닐 수 있는 연구 분야가 존재함을 의미함, 3) 날짜 유형(data type) 요소의 경우, 선택 및 불필요 요소로 응답한 사례가 존재. 이는 메타데이터 요소값 중 표준적 기술 방법 적용의 필요성을 인식하고 있지 못함을 의미, 4) 공헌자(contributor) 요소의 경우, 연구 분야에 따라 선택적 요소로 판단되지 않는 경우 발생, 5) 데이터 셋 범위(coverage)를 불필요 요소로 판단하는 연구 분야 상존, 6) 파일 크기(size), 파일 크기 단위(unit), 위치 정보(Geographic Location)는 선택요소 이거나 불필요 요소로 판단되는 경우가 존재하였다.

이상의 내용을 종합적으로 판단해 볼 때, 스키마 설계에 반영되어야 하는 시사점과 설계가 완료된 스키마 활용과 관련된 시사점이 도출되었다. 첫 째, 스키마 생성 시 분야별 특징을 포용할 수 있도록 확장성 있는 설계가 이루어져야 한다. 이는 특정 요소에 대해 필요정도가 다르기 때문이다. 둘째, 본 연구 결과로 도출된 메타데이터 스키마 활용 시에 연구자 교육이 반드시 필요하다 판단된다. 특히 연구 데이터 출판의 기대효과에 대해 연구자들의 이해를 돕도록 해야 한다. 한편, 한국한의학연구원에 구축예정인 데이터 리포지터리가 KISTI에서 운영하는 국가연구 데이터플랫폼과 연동이 되어야하기 때문에 도출된 요소를 한국과학기술정보연구원(KISTI)에서 정의한 연구 데이터 스키마를 대상으로 요소검토를 진행하였다.

〈표 4〉는 한의학 연구 데이터 분석 및 관련 표준 검토, 연구자 인터뷰를 통해 도출한 한국한의학연구원 메타데이터 사전설계 요소와 KISTI 메타데이터 요소(통합 공통 및 파일 메타요소)를 비교한 crosswalk를 보여준다. 분석된 특징은 다음과 같다. 첫째, 사전에 설계된 한의학연구원 메타데이터 요소는 개별적인 연구 데이터를 기술하기 위한 메타데이터 요소들로 구성되어 있다. 이와 반면에 KISTI 메타데이터 요소는 데이터 셋과 파일 수준의 메타데이터 요소를 별도로 구성하여 제시하고 있다. 이는 TTA 표준에서 제시한 것을 따르고 있다. 따라서 본 연구에서 사전에 설계한 메타데이터 요소를 데이터 셋과 파일 수준의 메타데이터 수준으로 구분할 필요가 있다. 이를 위해서, 사전 설계된 요소의 용도에 대한 구체적 재정의가 필요하다. 둘째, 보다 상세하게 도출된 메타데이터 요소를 수용하여 메타데이터 요소가 확장될 필요가 있다. 이는 한국한의학연구원에서 수집・생산되는 연구 데이터의 특징을 특화해 반영하고자 함을 목표로 한다. 또한 연구 데이터 거버넌스 체계 구축을 위해 KISTI가 수집하고자 하는 메타데이터 요소 항목에 부합하기 위함을 목적으로 한다. 셋째, 확장 가능한 메타데이터 요소를 설계하기 위해서 통제어로 제시 가능한 메타데이터 요소는 하나의 요소로 도출하며 그 값은 통제어로 관리할 수 있도록 해야 한다.

4.2 한국한의학연구원 메타데이터 스키마

한국한의학연구원 메타데이터 사전설계 요소와 KISTI 메타데이터 요소(통합 공통 및 파

www.kci.go.kr

〈표 4〉 사전설계요소 및 KISTI 메타데이터 요소 Crosswalk

No.	한국한의학연구원 사전설계 요소		KISTI 메타데이터 요소(통합 공통 및 파일 메타)	
	메타 데이터 요소	표시상수	메타데이터 요소	표시상수
1	Identifier	식별자	Identifier, IdentifierType	식별자 / 식별자유형
2	Title	제목	Title, TitleEng	제목(한글/영문)
3	Creator	생산자	Creator, CreatorEng	생산자(한글/영문)
4	Publisher	출판자	Publisher, PublisherEng	발행처(한글/영문)
5	PublicationYear	출판년도	PublicationYear	발행연도
6	Date	날짜	BaseYear, CreateDate, RegDate, ModifyDate, EmbargoDate	기준년도, 생성일, 등록일, 수정일, 엠바고날짜
6.1	DateType	날짜 유형	n/a	n/a
7	Description	데이터셋 설명	Description, DescriptionEng	설명(한글/영문)
8	Subject	주제	SubjectName, SubjectShema, SubjectID	주제분류명, 주제정보스키마, 주제분류코드
9	Contributor	공헌자	Contributor, ContributorEng	기여자(한글/영문)
9.1	ContributorType	공헌자 유형	ContributorType	기여자유형
10	Contact	연락처	ContactName, ContactNameEng, ContactAddress, ContactPhone, ContactEmail	담당자명(한글/영문), 주소, 전화, 이메일
11	Rights	저작권	Rights	저작권
12	Keyword	키워드	Keyword / KeywordEng	키워드(한글/영문)
13	AccessType	접근 유형	accessType	접근권한
14	AccessRestriction	접근 제한	n/a	n/a
15	ResourceType	데이터셋 유형	Type	유형
16	FundingReference	펀딩 참조정보	Funder, Funding	과제펀딩기관, 연구비
17	Coverage	데이터셋 범위	n/a	n/a
18	Format	파일 형식	n/a	n/a
19	Size	파일 크기	n/a	n/a
20	Unit	파일 크기 단위	n/a	n/a
21	GeographicLocation	위치 정보	CollectedFrom	수집위치
22	n/a	n/a	Language	언어
23	n/a	n/a	Version	버전
24	n/a	n/a	ScienceClass	과학기술표준분류명
25	n/a	n/a	Location	소재위치
26	n/a	n/a	LandingPageURL	상세정보 URL
27	n/a	n/a	ExtraAttribute	기타속성
28	n/a	n/a	FileIdentifier	파일식별자
29	n/a	n/a	FileIdentifierType	파일식별자유형
30	n/a	n/a	FileTitleKor / FileTitleEng	파일명(한글/영문)
31	n/a	n/a	FileCreatorKor, FileCreatorEng	파일생산자(한글/영문)
32	n/a	n/a	CreatorOrgName	파일생성자 소속기관
33	n/a	n/a	FilePublisher	파일발행처
34	n/a	n/a	FileCreateDate	파일생성일
35	n/a	n/a	FileDescription	파일설명
36	n/a	n/a	FileRights	파일저작권
37	n/a	n/a	FileKeyword, FileKeywordEng	파일키워드(한글/영문)

www.kci.go.kr

일 메타요소)를 비교한 crosswalk를 바탕으로 설계된 최종적인 메타데이터 스키마를 설계하였다. 연구 데이터를 식별하기 위한 Identifier 요소를 포함하여 총 6개 필수요소와 연구 데이터의 주제를 기술하는 Subject 요소를 포함하여 총 7개 권고요소, 연구 데이터 저작권 정보를 기술하기 위한 Rights 요소를 포함한 총 16개 임의요소를 설계하였다.

본 연구를 통해 제안되는 스키마는 크게 세 가지 특징을 지니고 있으며 그에 따른 의미를 갖는다. 첫째, 데이터를 출판하기 위한 DataCite 스키마를 준용하고 있다. DataCite에서 발표한 최신 스키마 버전 4.1을 준용하고 있다. 따라서 본 스키마를 활용하여 한국한의한연구원의 데이터가 관리될 경우, 연구 데이터의 글로벌 출판이 가능하다. 둘째, KISTI에서 제안하고 있는 '통합 공통 및 파일 메타요소'를 포함하고 있다. 이를 통해 한국한의학연구원의 메타데이터를 국가적인 연구 데이터 거버넌스 체제에 출판 할 수 있다. 국가연구 데이터 플래폼을 통해서 한국한의학연구원으로의 인바운드 링크가 유입되어 연구원 데이터의 대국민 서비스가 가능하다. 셋째, 한국한의학연구원에서 연구 데이터 관리를 위해 기존에 사용하던 메타데이터 요소를 모두 포함하고 있으므로 본 연구에서 제안하는 메타데이터 요소를 기반으로 한 새로운 시스템 구축 시 성공적인 데이터 이관을 보장한다.

최종 설계된 한국한의학연구원 연구 데이터 관리 및 서비스를 위한 메타데이터 스키마는 〈표 5〉와 같다. 표에서 제시한 정의 및 허용 값으로 각 요소에 대한 이해가 가능하기 때문에 상세한 설명은 생략한다. 각 요소에는 해당 요소를 도출한 출처가 명시되어 있다. 이를 통해 해당 요소에의 원 정의를 추적할 수 있다. 이는 표준 요소 재사용을 의미하며, 향후, 본 스키마를 활용한 다양한 응용프로파일이 작성 시, 명확한 요소사용을 권고하기 위함이다.

〈표 5〉 한국한의학연구원 연구 데이터 관리 및 서비스를 위한 메타데이터 스키마

ID	메타 데이터 요소	표시상수	빈도	필수여부	정의, 허용 값, 예제, 제한 조건
1	Identifier	식별자	1	M	▫ DataCite 출판요소 ▫ 데이터셋의식별자또는식별값 ▫ 공식적인식별체계에맞는문자열이나번호를사용하는방법을권고함. ▫ 공식적인식별체계에제한이없으나, URI(웹주소인URL을포함), 디지털 객체 식별기호(DOI), 국가표준 디지털콘텐츠식별체계(UCI) 등이 포함됨.
1.1	identifierType	식별자유형	1	M	▫ DataCite 출판요소 ▫ 데이터셋〈Identifier〉요소의유형 ▫ '부록Ⅱ-1통제어:IdentifierType을참조하여해당값을선택함 ▫ 데이터셋을웹에공개또는출판하기위해반드시기술되어야함
1.2	altenateIdentifier	대체식별자	0-n	O	▫ 대체 식별자 ▫ 시스템 로칼 식별자 등을 기술함 ▫ Identifier에서 가리키는 동일한 데이터를 가리킴 ▫ 동일한 파일과 동일한 연구 데이터 위치를 가리킴
1.2.1	altenateIdentifer Type	대체식별자 유형	1	O	▫ 대체 식별자 유형 ▫ 로컬 식별자로서 접근번호 등을 사용함

www.kci.go.kr

ID	메타 데이터 요소	표시상수	빈도	필수여부	정의, 허용 값, 예제, 제한 조건
2	Title	제목	1-n	M	▫DataCite 출판요소 ▫데이터셋의이름또는명칭 ▫기술되는값의언어정보가함께기술되는것을권고함
2.1	titleLang	제목언어	1	M	▫KISTI 연계 요소 ▫데이터셋명칭을기술한언어 ▫Languagecode: ISO 639-3 사용
3	Creator	생산자	1-n	M	▫DataCite 출판요소 ▫데이터셋을생성또는묶은사람, 조직등을기술함 ▫기술되는값의언어정보가함께기술되는것을권고함 ▫데이터셋을웹에공개또는출판하기위해반드시기술되어야함
3.1	creatorLang	생산자 기술 언어	1	M	▫KISTI 연계 요소 ▫데이터셋생성자를기술한언어 ▫Languagecode: ISO 639-3 사용
4	Publisher	출판자	1	M	▫DataCite 출판요소 ▫데이터셋을웹에공개또는출판하는주체 ▫연구자,기관,서비스명등을기술함
4.1	publisherLang		1	M	▫KISTI 연계 요소 ▫출판사명칭을기술한언어
5	PublicationYear	출판년도	1	M	▫DataCite 출판요소 ▫데이터셋이웹상에서접근또는이용가능해진연도 ▫데이터셋의공개범위가 'Embargoed(이용지연)'으로설정되면,이용지연이종료되는연도를기술함 ▫YYYY형식의4자리연도로기술함
6	ResourceType	데이터셋 유형	1	M	▫DataCite 출판요소 ▫자원의상세유형을기술함 예시: Dataset/CensusData,Text/ConferenceAbstract와같이 '일반적자원유형/상세자원유형' 형식으로기술
6.1	resourceTypeGeneral	데이터셋 일반유형	1	M	▫데이터셋의 일반적 자원유형을 기술함 ▫TTA 부록 II-통제어:ResourceTypeGeneral을참조하여해당값을선택함 Audiovisual,Collection,DataPaper,Dataset,Event,Image,InteractiveResource,Model,PhysicalObject,Service,Software,Sound,Text,Workflow,Other
7	Date	날짜	0-n	R	▫데이터셋이 생애주기 동안 발생한 다양한 이벤트 날짜 ▫기술되는 값은 ISO 8601[W3CDTF]에 정의된 방식 중 'YYYY-MM-DDThh:mm:ssTZD' 형식을 사용할 것을 권고함 ▫예: 1997-07-16T19:20:30+01:00
7.1	dateType	날짜 유형	1	M	▫데이터셋 〈Date〉 요소의 유형 ▫TTA 부록 II-통제어:DatasetDateType을참조하여해당값을선택함 ▫Accepted,Available,Copyrighted,Collected,Created,Issued,Submitted,Updated,Valid,BaseYear(기준년도),CreateDate(생산일),RegDate(등록일),ModifyDate(수정일),EmbargoDate(엠바고일)
8	Description	데이터셋 설명	0-n	R	▫데이터셋의 내용에 대한 설명 ▫기술방법에는제한이없으나주로맥락정보를상세히기술 ▫기술되는값의언어정보가함께기술되는것을권고함
8.1	descriptionLang	데이터셋 설명 언어	1	M	▫KISTI 연계 요소 ▫데이터셋설명을기술한언어 ▫Languagecode: ISO 639-3 사용
8.2	descriptionType	데이터셋 설명 유형	1	M	▫DataCite 연계 요소 ▫데이터셋설명의유형을기술함 ▫TTA 부록 II-통제어:DescriptionType을참조하여해당값을선택함 Abstract,Methods,SeriesInformation,TableOfContents,TechnicalInfo,Other

www.kci.go.kr

ID	메타 데이터 요소	표시상수	빈도	필수여부	정의, 허용 값, 예제, 제한 조건
9	Subject	주제	0-n	R	▫본 연구에서 도출된 주제분류를 코드화해 사용 가능 ▫데이터셋 주제 정보를 기술하기 위해 하위요소를 묶는 요소 ▫예:XML사례 〈SubjectSubjectScheme="DFG"〉 〈SubjectID〉31702〈/SubjectID〉 〈SubjectName〉HumanGeography〈/SubjectName〉〈/Subject〉 〈SubjectSubjectScheme="DFG"〉 〈SubjectID〉31〈/SubjectID〉 〈SubjectName〉Chemistry〈/SubjectName〉〈/Subject〉 ▫예:JSON사례 {Subject: [{"SubjectScheme":"DFG","SubjectID":"31702","SubjectName":"HumanGeography"}, {"SubjectScheme":"DFG","SubjectID":"31","SubjectName":"Chemistry"}]}
9.1	subjectID	주제분류코드	1	R	▫주제분류표에서 제시되는 id(주제코드 값)를 기술함 ▫DFG주제분류명으로HumanGeography을사용하는경우:31702 ▫Subject요소활용예시참조
9.2	subjectSheme	주제정보 스키마	1	R	▫데이터셋 주제정보 기술을 위해 참고 되는 정보원을 기술함 ▫웹으로접근가능할경우웹주소를기술할수있으며,정보원의명칭을기술할수있음 ▫DFG주제분류를사용하는경우:DFG ▫Subject요소활용예시참조
9.3	subjectName	주제분류명	1	M	▫주제분류표에서 제시되는 주제 명칭을 기술함 ▫DFG주제분류코드로31702를사용하는경우:HumanGeography ▫Subject요소활용예시참조
10	Contributor	공헌자	0-n	O	▫데이터셋의 수집, 생산, 예산지원 등에 관계된 사람이나 기관 등을 기술함 ▫EPSRC-EngineeringandPhysicalSciencesResearchCouncil
10.1	contributorLang	공헌자 기술 언어	1	M	▫공헌자 혹은 공헌기관 명칭을 기술한 언어 ▫Languagecode: ISO 639-3 사용
10.2	contributorType	공헌자 유형	1	R	▫데이터셋의 수집, 생산, 예산지원 등에 관계된 사람이나 기관 등의 유형 ▫'TTA 부록 II-통제어:ContributorType을참조하여해당값을선택함
11	Contact	연락처	0-n	O	▫연락처 정보를 기술하기 위해 하위요소를 묶는 요소 ▫데이터셋에대한문의사항에대응할수있는담당자의연락처정보를기술함 ▫이메일,전화,웹주소등을값으로기술할수있음
11.1	contactName	담당자명	1	M	▫해당 과제/프로젝트와 관련된 담당자를 기술함
11.2	contactNameLang	담당자명 기술 언어	1	M	▫해당 과제/프로젝트와 관련된 담당자를 기술한 언어 ▫Languagecode: ISO 639-3 사용
11.3	contactAddress	담당자 주소	0-1	O	▫해당 과제/프로젝트와 관련된 담당자의 주소
11.4	contactPhone	담당자 전화	0-1	O	▫해당 과제/프로젝트와 관련된 담당자의 유선전화번호, 핸드폰번호 (예: 0428690652)
11.5	contactEmail	담당자 이메일	0-1	O	▫해당 과제/프로젝트와 관련된 담당자의 이메일 (예: kildong@korea.kr)
12	Rights	저작권	0-1	R	▫데이터셋의 저작권 정보를 기술함 ▫예시:해당데이터셋의저작권은연구자에게귀속됨
13	Keyword	키워드	0-n	R	▫데이터셋의 내용이 지닌 주제를 자연어로 기술함 ▫데이터셋에대한키워드검색에노출시킬용어를기술함 ▫예시:DNA,nucleosomepositioning
13.1	keywordLang	키워드 기술 언어	1	M	▫키워드를 기술한 언어 ▫Languagecode: ISO 639-3 사용

www.kci.go.kr

ID	메타 데이터 요소	표시상수	빈도	필수여부	정의, 허용 값, 예제, 제한 조건
14	FileFieldKeyword	키워드	0-n	R	▫메타데이터 파일의 필드 설명 정보 중 검색에 노출시킬 용어를 기술함 ▫예시: 체중, 전위
14.1	fileFieldKeyword Lang	키워드 기술 언어	1	M	▫키워드를 기술한 언어 ▫Languagecode: ISO 639-3 사용
15	AccessType	접근 유형	0-1	O	▫데이터셋의 접근 또는 이용 유형 ▫TTA 부록 II-통제어:DatasetAccessType을참조하여해당값을선택함
16	AccessRestriction	접근 제한	0-1	O	▫데이터셋의 접근제한 또는 이용제한 유형 ▫TTA 부록 II-통제어:DatasetAccessRestrictionType을참조하여해당값을선택함
17	FundingReference	펀딩 참조정보	0-n	O	▫펀딩 정보를 기술하기 위해 하위요소를 묶는 요소
17.1	funderName	펀딩기관	1	M	▫KISTI Funder 요소 연계 ▫펀딩하는기관(예:과학기술정보통신부)
17.2	funderIdentifier	펀딩기관 식별자	0-1	O	▫펀딩기관 식별자를 기술함 ▫예시:https://doi.org/10.13039/100
17.2.1	funderIdentifierType	펀딩기관 식별자 유형	0-1	O	▫펀딩기관 식별자 유형을 기술함 ▫ISNI,GRID,CrossrefFunder,한국연구재단,Other
17.3	awardNumber	과제번호	0-1	M	▫과제번호를 기술함 ▫예시: GBMF3859.01
17.4	awardURI	과제URI	0-1	O	▫과제 정보가 기술된 웹 주소를 기술함 ▫예시: https://www.moore.org/grants/list/
17.5	awardTItle	과제명	1	M	▫과제명을 기술함 ▫예시: SocioenvironmentalMonitoringoftheAmazonBasinandXingu
17.6	awardBudget	과제예산	0-1	O	▫KISTI Funding 요소 연계 ▫과제예산을자유롭게기술함
18	Coverage	데이터셋 범위			▫특정 시기는 [W3CDTF] - the "Date/Time Format" of the W3C 형식으로 기입함 - 연도: CCYY (e.g. 1999) - 연도, 월: CCYY-MM (e.g. 1999-07) - 연도, 월, 일: CCYY-MM-DD (e.g. 1999-07-01) - 연도, 월, 일, 시, 분, 그리니치 표준시: CCYY-MM-DDThh:mmTZD (예) 1999-07-01T19:30+10:00 - 연도, 월, 일, 시, 분, 초, 그리니치 표준시: CCYY-MM-DDThh:mm:ss TZD (e.g. 1999-07-01T19:30:45+10:00) ▫기간은시작일과종료일로기입함 - 1999 to 2000, 1999-07-01 to 1999-07-30 ▫수집지역은지역명을자유기술식으로기입함 - 서울시 (수집지역)
19	GeographicLocation	위치 정보	0-n	O	▫KISTI CollectedFrom 요소 연계 ▫DataCiteMetadataSchema4.1요소중geoLocationPoint(점경도,점위도),geoLocationBox(서쪽경계경도,동쪽경계경도,남쪽경계위도,북쪽경계위도),geoLocationPlace(지역이름이나지역에대한설명),geoLocationPolygon(다각형지역)를선택적으로기술하거나자유로운텍스트로기술함
20	Language	언어	0-1	O	▫데이터셋을 기술한 언어 ▫Languagecode: ISO 639-3 사용
21	Version	버전	0-1	O	▫KISTI 요소 연계 ▫데이터셋버전
22	ScienceClass	과학기술표준 분류명	0-n	R	▫데이터셋의 과학기술표준분류명을 기술함

www.kci.go.kr

ID	메타 데이터 요소	표시상수	빈도	필수여부	정의, 허용 값, 예제, 제한 조건
23	Location	소재위치	0-n	O	▫ 데이터셋의 물리적/논리적 접근 위치를 기술함 ▫ 물리적인주소가될수있으며,논리적인파일경로,웹주소가될수있음
24	LandingPageURL	상세정보 URL	0-1	O	▫ 데이터셋 정보를 제공하는 URL
25	ExtraAttribute	기타속성	0-1	O	▫ 데이터셋에 특화된 속성들을 기술함. 원 데이터에 대한 속성 설명 등을 기술함
26	RelatedIdentifier	관련자원 식별자	0-n	R	▫ 연구 데이터와관련된자원식별자를기술함 ▫ 글로벌식별자를기술해야함
26.1	relatedIdentifier Type	관련자원 식별자 유형	1	M	▫ 관련된자원의식별자유형을기술함 ▫ 부록-통제어:relatedIdentifierType을참조하여해당값을선택함 ▫ 예시: ARK,arXiv,bibcode,DOI,EAN,EISSN,Handle,IGSN,ISBN,ISSN,ISTC,LISSN,LSID,PMID,PURL,UPC,URL,URN
26.2	relationType	관련자원 관계 유형	1	M	▫ 관련된자원의유형을기술함 ▫ 부록-통제어:relationType을참조하여해당값을선택함 ▫ 예시:IsCitedBy,Cites,IsSupplementTo,IsSupplementedBy,IsContinuedBy,Continues,IsDescribedBy,Describes,HasMetadata,IsMetadataFor,HasVersion,IsVersionOf,IsNewVersionOf,IsPreviousVersionOf,IsPartOf,HasPart,IsReferencedBy,References,IsDocumentedBy,Documents,IsCompiledBy,Compiles,IsVariantFormOf,IsOriginalFormOf,IsIdenticalTo,IsReviewedBy,Reviews,IsDerivedFrom,IsSourceOf,IsRequiredBy,Requires
26.3	relatedResource Title	관련자원 제목	1	R	▫ 관련된 자원의 제목을 기술함
27	fileIdentifier	파일식별자	0-n	O	▫ 파일 식별자를 기술함 ▫ 파일을 유일하게 식별하기 위한 URI 등을 사용 할 수 있음
27.1	fileIdentifierType	파일식별자 유형	0-1	O	▫ 파일 식별자 요소가 사용되었을 경우, 필수 요소로 기술되어야 함
27.2	fileTitle	파일명	1-n	M	▫ 해당 파일의 제목
27.2.1	fileTitleLang	파일명 기술 언어	1	M	▫ 해당 파일의 제목을 기술한 언어
27.3	fileCreator	파일생산자	0-n	R	▫ 원 데이터 생산자
27.3.1	fileCreatorLang	파일생산자 기술 언어	1	M	▫ 원 데이터 생산자를 기술한 언어
27.3.2	creatorOrgName	파일생성자 소속기관	0-1	O	▫ 생성자 소속기관
27.4	filePublisher	파일발행처	0-1	O	▫ 파일을 웹에 공개 또는 출판하는 주체
27.5	fileCreateDate	파일생성일	0-1	O	▫ 파일을 생성한 날짜 ▫ 예시: 20180101
27.6	fileDescription	파일설명	0-1	O	▫ 파일 내용에 대한 설명
27.7	fileRights	파일저작권	0-1	O	▫ 파일의 저작권 정보
27.8	fileKeyword	파일키워드	0-1	O	▫ 파일 내용이 의미를 명사형으로 기술
27.8.1	fileKeywordLang	파일 키워드 기술 언어	0-1	O	▫ 파일 내용이 의미를 명사형으로 기술한 언어
27.9	fileFormat	파일 형식	0-1	O	▫ 파일 형식을 확장자명과 간단한 설명으로 기입함 ▫ 예시: PPT (Power Point Template))
27.A	fileSize	파일 크기	0-1	O	▫ 파일의 크기(용량)를 단위 없이 값만 기술함 ▫ 바이트단위로입력하는것을권고함
27.A.1	fileSizeUnit	파일 크기 단위	1	M	▫ 파일 〈Size〉 요소의 단위를 기술함 ▫ TTA 부록 II-통제어:FileSizeUnitType을참조하여해당값을선택함 ▫ 단위가MB인경우:MegaByte ▫ 단위가바이트인경우:Byte

www.kci.go.kr

4.3 메타데이터 요소 스키마 검증

메타데이터 요소 스키마 검증을 위해, 종료과제 연구 데이터 제출 시스템과 이미 사용 중이던 메타데이터 항목을 대상으로 crosswalk를 작성하였다. 한의학 연구원에는 종료과제에서 산출되는 연구 데이터를 수집하기 위한 시스템을 개발하여 운영 중이다. 이 시스템에서 사용하는 메타데이터 요소는 〈표 6〉의 기존 시스템 관리요소에 나와 있는 것처럼 데이터 식별자, 데이터 등록자 및 참여자 정보, 관련 과제 정보, 데이터에 대한 설명정보에 해당하는 것들로 구성되어 있다. 종료과제 연구 데이터 제출 시스템에서 사용하는 요소를 검토한 결과 '연구책임자', '수행부서', '연구 데이터 제출 가능여부', '연구 데이터 미제출 사유', '연구 데이터 수정요청 사유', '제출가능 여부', '데이터 파일 미제출 사유' 요소를 제외한 모든 요소를 설계요소가 수

〈표 6〉 기존 시스템과의 비교

기존 시스템의 관리요소	대응 한의학 연구 데이터 메타데이터 요소	기존 시스템의 관리요소	대응 한의학 연구 데이터 메타데이터 요소
데이터저장관리번호(DB key)	Identifier identifierType	자료유형	ResourceType resourceTypeGeneral
등록(신청자)	contactName	라이선스	Rights
전화번호	Contact	데이터공개여부	AccessType
핸드폰	contactPhone	데이터비공개 사유	AccessRestriction
과제정보(제출과제)	awardNumber	키워드	Keyword keywordLang
과제명	awardTItle	생성(생산)일자	Date dateType
연구책임자	n/a	실험기기	ExtraAttribute
수행부서	n/a	자료처리 SW 및 버전	ExtraAttribute
과제기간	Coverage	요약설명	Description
과제예산	awardBudget	참여자 정보	Creator creatorLang
연구 데이터 제출 가능여부	n/a	관련 논문	RelatedIdentifier relatedIdentifierType relationType relatedResourceTitle
연구 데이터 미제출사유	n/a	데이터파일 제출가능여부	n/a
연구 데이터 수정요청사유	n/a	데이터 포맷	fileIdentifierType
데이터명칭(한글) 데이터명칭(영문)	Title titleLang	데이터파일 미제출 사유	n/a
데이터분류 연구방법분류 질환/병증 분류	Subject subjectID subjectScheme subjectName	데이터 첨부파일 메타데이터 첨부파일	FileIdentifier fileIdentifierType fileTitle fileCreator fileCreateDate

www.kci.go.kr

용 가능하였다. '연구 책임자'와 '수행부서' 정보는 설계 된 'Contributor'와 'ContributorType' 요소를 활용해 기술 가능하다. 그 외 요소는 종료과제에 대한 연구 데이터 제출 시스템에서 별도 관리해야 하는 항목으로 판단하여 본 연구에서 제안하는 메타데이터 요소에는 포함하지 않았다.

〈표 6〉은 종료 수행과제 연구 데이터를 수집하기 위해 개발했던 시스템의 관리 요소들에 대응되는 한의학 메타데이터 요소를 보여준다. 대부분의 관리요소들은 한의학 메타데이터 요소로 기술이 가능함을 확인하였다. 〈표 6〉에는 메타데이터 요소와 대응되지 않는 일부 시스템 관리 요소들도 있다. '연구책임자', '수행부서'는 시스템 내부 관리 항목이고 메타데이터로서 관리될 만한 요소는 아니다. '연구 데이터 제출 가능여부', '연구 데이터 미제출사유', '연구 데이터 수정요청사유', '데이터파일 제출가능여부', '데이터파일 미제출사유'는 연구 데이터 수집 과정 중에 필요한 정보들을 담고 있는 요소들로서 마찬가지로 메타데이터 요소로서 추후에 데이터의 활용을 위해서 이용될 만한 것들이 아니다. 따라서 제안된 연구 데이터 메타데이터 요소들만으로 기존 시스템에서 사용된 요소들을 대체할 수 있음을 확인하였다. 본 연구에서 개발된 메타데이터 스키마는 향후에 기관 연구 데이터 리포지터리 시스템에 XML 인코딩 스키마 형태로 반영되어 활용할 수 있을 것이다. 실제 운영을 바탕으로 적용사례 연구를 수행하여 더욱 효율적인 스키마로 개선 될 수 있을 것이다.

5. 결론 및 제언

본 연구를 통해 한국한의학연구원에서 생산되는 연구 데이터를 체계적으로 관리하고 재사용하기 위한 메타데이터 요소를 설계하였다. 최상위 요소로서 Identifier를 포함한 27개 요소가 도출되었으며, Title(제목)과 같은 필수요소 29개, Description(설명)과 같은 권고요소 13개, Alternate Identifier(대체식별자)와 같은 선택요소 31개를 도출하였다. 한국한의학연구원에서 운영중인 기존 시스템의 메타데이터 관리요소와 본 연구에서 도출한 메타데이터 요소를 비교하여 기존 요소에서 누락되는 요소가 없음을 확인 하였다. 한의학 분야 주제분류를 제안했으며, 제안된 주제분류를 코드화해 스키마에 적용할 수 있을 것이다.

본 연구에서는 향후에 개발될 한국한의학연구원 데이터 리포지터리와 한국과학기술정보연구원에서 구축중인 국가연구 데이터플랫폼과의 연계를 위한 메타데이터 요소도 검토, 적용하였다. 따라서 미래지향적이며 확장성 있는 메타데이터 스키마로 판단된다.

데이터중심 R&D 활성화와 데이터 분석을 활용한 다분야 융합・공동연구가 활성화됨에 따라 연구 데이터 공유・활용 체계의 필요성 부각되는 국내외의 패러다임의 변화에 부응하여 한의학 연구분야 연구 데이터를 효율적으로 수집・관리・공유・활용하기 위해서 연구 분야별 전문가 자문을 통해서 각 분야에서 생성되는 연구 데이터의 특성을 파악하고 이를 명확하게 구분・관리할 수 있는 분류체계를 기존의 분류체계와 선행연구를 바탕으로 고안하였다. 세 가지 상이한 분류체계를 통해서 연구자

www.kci.go.kr

가 자신의 연구 데이터를 좀 더 명확한 태깅을 통해서, 향후 관리의 용이성 및 검색의 정확성을 향상할 수 있을 것으로 기대된다. 한의학 연구 데이터 메타데이터 요소들은 단일 스키마로 이루어져 다양한 한의학 연구 데이터를 포괄할 수 있는 범용성과 유연성을 가지도록 설계되었다. 그러나 데이터는 그 형식에 있어서 제한이 없으므로 한정된 메타데이터 요소들만으로 데이터와 관련한 모든 사항을 완벽하게 기술하는 것이 용이하지 않은 경우도 적지 않다. 예를 들어 유전체 염기서열 데이터와 같이 특정 실험 기기로 부터 산출되는 데이터의 경우 데이터를 산출하는데 사용된 기기 및 프로그램의 구체적인 설정 값과 같은 것 역시 주요한 메타데이터 정보로 활용될 수 있다. 연구 데이터에 대한 더욱 상세한 사항들은 Readme 파일과 같은 보조 파일에 자연어로 기술하여 관리하는 것도 연구 데이터가 효율적으로 활용되기 위한 보완책이 될 수 있을 것이다(Lee and Stvilia 2017). 본 연구를 통해 제안되는 메타데이터 스키마는 한국한의학연구원에서 수집, 생산되는 데이터와 정보, 지식을 관리하고 검색하는데 사용되기를 바란다.

참 고 문 헌

[1] 과학기술정보통신부. “국가과학기술표준분류.” 2018 [cited 2019. 1. 6.]

[2] 김선태 외. 2017.『연구 데이터 관리 및 공유를 위한 메타데이터(The metadata for the managing and sharing research data)』. 한국정보통신기술협회. 정보통신단체표준 TTAK.K0-10.0976.

[3] 신은정. 2018. “국가R&D사업 데이터 관리 계획 도입의 방향.”『STEPI 과기정책포럼』.

[4] 조성복. 2012. “한의학 자원 DB 통합을 위한 정보화 전략 기획(ISP)에 관한 연구.”

[5] 한국보건산업진흥원. “보건산업기술분류체계.” 2013 [cited 2019. 1. 6.]

[6] 한국보건산업진흥원. “HT융합기술의 분류체계 연구.” 2015 [cited 2019. 1. 6.]

[7] 한국보건산업진흥원. “HT표준기술분류체계.” 2014 [cited 2019. 1. 6.]

[8] Adams Becker, S. et al. 2017. NMC Horizon Report: 2017 Library Edition. Austin, Texas: The New Media Consortium.

[9] DataCite Statistics. [online] [cited 2018. 12. 11.] 〈https://stats.datacite.org/#tab-datacentres〉

[10] Europe 2020 strategy. [online] [cited 2019. 3. 6.] 〈https://bit.ly/2gEXPR2〉

[11] European Commission. 2016. Open Innovation, Open Science, Open to the World. ISBN 978-92-79-57346-0 doi:10.2777/061652

[12] Horizon 2020. [online] [cited 2019. 2. 6.] 〈https://bit.ly/1Kz2D4x〉

www.kci.go.kr

[13] Innovation Union. [online] [cited 2019. 2. 6.] 〈https://bit.ly/2D1fRKN〉

[14] Johnson, R., Chiarelli, A. and Parsons, T. 2016. Data asset framework (DAF) survey results. figshare. Fileset. https://doi.org/10.6084/m9.figshare.3796305.v4

[15] JST. 2017a. JST Policy on Open Access to Research Publications and Research Data. Management. https://www.jst.go.jp/EN/about/openscience/policy_openscience_en.pdf

[16] JST. 2017b. Implementation Guidelines: JST Policy on Open Access to Research Publications and Research Data Management. https://www.jst.go.jp/EN/about/openscience/guideline_openscience_en.pdf

[17] National Academy of Sciences (US), National Academy of Engineering (US) and Institute of Medicine (US) Committee. 2009. Ensuring the Utility and Integrity of Research Data in a Digital Age. Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age. Washington (DC): National Academies Press (US). https://www.ncbi.nlm.nih.gov/books/NBK215264/ doi: 10.17226/12615

[18] Röhrig, B. et al. 2009. "Types of study in medical research: part 3 of a series on evaluation of scientific publications." *Deutsches Arzteblatt International*, 106(15): 262.

[19] Rolando, L. 2015. "Data Management Plans as a Research Tool." *Bul. Am. Soc. Info. Sci. Tech*, 41: 43-45. doi:10.1002/bult.2015.1720410510

[20] Rylance, R. et al. 2016. Concordat on Open Research Data. https://www.ukri.org/files/legacy/documents/concordatonopenresearchdata-pdf/

[21] SPEC Kit 334. [online] [cited 2019. 2. 6.] 〈http://publications.arl.org/Research-Data-Management-Services-SPEC-Kit-334/〉

[22] Starr, J. and Gastl, A. 2011. "isCitedBy: A Metadata Scheme for DataCite." *D-Lib Magazine*, 17(1).

[23] Süt, N. 2014. "Study designs in medicine." *Balkan medical journal*, 31(4): 273.

[24] Tenopir, C. et al. 2017. "Research Data Services in European Academic Research Libraries." *LIBER Quarterly*, 27(1): 23-44. DOI:http://doi.org/10.18352/lq.10180

[25] Uman, L. S. 2011. "Systematic reviews and meta-analyses." *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 57.

[26] Zeng, X. et al. 2015. "The methodological quality assessment tools for preclinical and clinical studies, systematic review and meta-analysis, and clinical practice guideline: a systematic review." *Journal of evidence-based medicine*, 8(1): 2-10.

www.kci.go.kr

• **국문 참고자료의 영어 표기**
(English translation / romanization of references originally written in Korean)

[2] Kim, Suntae et al. 2017. "The metadata for the managing and sharing research data." Telecommunications Technology Association. TTAK.K0-10.0976

[3] Shin, Eun-jung. 2018. "Introduction of the National R&D Project Data Management Plan" *STEPI Science and Technology Policy Forum.*

[4] Cho, SungBok. 2012. "A Study on Strategy for Oriental Medicine resources An Integrated Database System."

www.kci.go.kr

[부록 1] 한의학 연구 데이터 분류

• 진단 연구 데이터 분류

대분류	중분류	소분류
진단 연구 데이터	사진(四診)	지표/설문
		재현성
		유효성/정확성
		기타
	변증진단	지표/설문
		재현성
		유효성/정확성
		기타
	체질진단	지표/설문
		재현성
		유효성/정확성
		기타
	진단기기	생체지표
		유효성/정확성
		안전성
		시판후평가
		기타
	미분류	

• 약물 치료 연구 데이터 분류

대분류	중분류	소분류
약물 치료 연구 데이터	한약자원	기원/감별
		구성성분
		약리기전
		유효성
		안전성
		생산/가공
		포장/보관/유통
		복용/투약
		기타
	한약방제	구성성분
		약리기전
		유효성
		안전성
		가공/포장/보관
		복용/투약
		기타
	한약제품	후보소재
		유효성
		안전성
		복용/투약
		시판후평가
		기타
	미분류	

• 비약물 치료 연구 데이터 분류

대분류	중분류	소분류
비약물 치료 연구 데이터	침/뜸/부항	치료기전
		유효성
		안전성
		기타
	추나/도인	치료기전
		유효성
		안전성
		기타
	경락경혈	작용기전
		기타
	치료기기	생체지표
		유효성/정확성
		안전성
		시판후평가
		기타
	미분류	

• 예방/협진 연구 데이터 분류

대분류	중분류	소분류
예방/협진 연구 데이터	예방/예측	유효성
		안전성
		예측모형
		기타
	재활치료	치료기전
		유효성
		안전성
		기타
	양한방협진	진료모형
		유효성
		안전성
		효율성
		기타
	미분류	

www.kci.go.kr

• 정보 연구 데이터 분류

대분류	중분류	소분류
정보 연구 데이터	문헌정보	디지털화
		추출/가공
		번역
		기타
	생물정보	시퀀스
		발현/형질
		구조/기능
		네트워크/시스템
		기타
	의료정보	표준
		보안/암호
		모델
		공유
		라이프로그
		기타
	미분류	-
외부/오픈 데이터		
미분류		

• 정책/교육 연구 데이터 분류

대분류	중분류	소분류
정책/ 교육 연구 데이터	정책	법령
		정책
		백서
		연구동향
	통계	기술동향
		산업동향
		기타
		통계연감
		산업실태조사
		기타
	표준	단체표준
		국가표준
		국제표준
		기타
	교육	역량평가
		문제기반학습
		e-Learning
		기타
	미분류	

www.kci.go.kr