

OAK 메타데이터 요소 분석 및 수정(안) 제안에 관한 연구*

A Study on Analysis and Modification of OAK Metadata Elements

노 지 현(Jee-Hyun Rho)**

이 은 주(Eun-Ju Lee)***

이 미 화(Mihwa Lee)****

< 목 차 >

I. 서론	1. OAK 메타데이터 수정(안)의 도출 근거
II. OAK 참여기관의 메타데이터 요소 분석	2. OAK 메타데이터 수정(안)
III. 국내외 기관 리포지터리의 메타데이터 사례	V. 결론
IV. OAK 메타데이터 수정(안) 제안	

초 록

OAK(Open Access Korea)는 오픈 액세스 지식정보의 공동 활용 및 통합서비스 체제 구축을 위해 국가적 차원에서 추진하고 있는 기관 리포지터리 사업이다. 그러나 메타데이터의 구축과 공유에 있어 다양한 문제를 내재하고 있어 OAK 메타데이터 요소에 대한 개선이 시급한 상황이다. 이에 본 연구에서는 OAK 참여기관에서 사용하는 메타데이터 요소를 세밀하게 분석하고, OAK와 유사한 성격의 국내외 리포지터리 사례를 참조해 가면서, OAK 메타데이터 수정(안)을 제안하였다. 구체적으로, OAK 참여기관의 메타데이터 현황을 조사하기 위해 17개 기관의 사례를 입수하여 분석하였으며, 참여기관에서 사용하는 메타데이터 요소의 용도가 명확하지 않은 경우 기관 리포지터리 담당자와의 인터뷰를 통해 보완하였다. 더불어, DSpace, EPrints, BEPress, ETD-db, dCollection 리포지터리의 메타데이터 사례를 분석한 다음, OAK 참여기관 실무자들과의 대면 회의 및 이메일 교환을 통해 최종적인 OAK 수정(안)을 도출하였다.

키워드: OAK, 기관 리포지터리, 메타데이터, 더블린코어, DSpace, 응용프로파일

ABSTRACT

OAK(Open Access Korea) is a national repository project to support free access to Korea's open access knowledge information. However various problems and limitations are recognized when organizing and sharing metadata with OAK-IR system. This study aims to suggest a modified metadata standard to improve the quality of OAK metadata. To the end, this study analyzed all metadata elements used by OAK participating organizations and compared the metadata with the metadata of other typical repository systems. In details, metadata organized in 17 OAK participating organizations were analyzed with their metadata specifications, input data examples, and practitioner interviews. And a case study on DSpace, EPrints, BEPress, ETD-db, dCollection was conducted. The final modified OAK metadata elements were suggested through group discussion and coordination by the practitioners of OAK participating organizations, OAK supervisors, and the developers of OAK-IR system.

Keywords: OAK, Open Access Korea, Institutional Repository, Metadata, Dublin Core, DCMES, DSpace, EPrints, BEPress, ETD-db, SOMS, Application profile

* 본 연구는 2016년 국립중앙도서관 『OAK 메타데이터 요소 분석 및 입력지침 개발』의 일부를 바탕으로 수정 보완한 것임.

** 부산대학교 문헌정보학과 교수(jhrho@pusan.ac.kr) (제1저자)

*** 동의대학교 문헌정보학과 조교수(ejlee@deu.ac.kr) (공동저자)

**** 공주대학교 문헌정보교육과 부교수(leemh@kongju.ac.kr) (교신저자)

•논문접수: 2017년 2월 19일 •최초심사: 2017년 2월 28일 •게재확정: 2017년 3월 17일

•한국도서관·정보학회지 48(1), 137-160, 2017. [http://dx.doi.org/10.16981/kliiss.48.201703.137]

I. 서론

OAK(Open Access Korea)는 오픈 액세스 지식정보의 공동 활용 및 통합서비스 체제 구축을 목적으로 대학, 공공기관, 연구소, 기업체 등이 참여하는 지식협력체이다. 2009년 문화체육관광부의 『국가도서관 지식콘텐츠의 창조적 관리 및 확산 사업』으로 시작하였으며, 2014년 국립중앙도서관에서 인수하여 『OAK 국가 지식정보 구축 및 확산 사업』으로 추진하고 있다. OAK 사업을 보다 효과적으로 추진하기 위해서는 개별 기관에서 생산한 연구성과물을 체계적으로 관리할 수 있는 표준화된 소프트웨어의 개발 및 보급이 선결되어야 한다. 이에 OAK에서는 MIT에서 개발한 DSpace를 참조하여 OAK-IR(Institutional Repository)을 개발한 후 참여기관들이 공동으로 활용할 수 있도록 보급하고 있다.

2017년 2월 현재 OAK-IR을 이용하는 기관 리포지터리는 총 31개이다. 이를 기관의 성격에 따라 구분하면, 대학 10개, 연구소 11개, 공공기관 6개(지역대표도서관 2개 포함), 의학도서관 4개이며, 이들 리포지터리에 등록된 지식정보는 학위논문, 학술지논문, 보고서, 특허, 단행본, 멀티미디어자료 등 총 55만 여 건에 이르고 있다. OAK에서는 OAK-IR에 구축된 지식정보에 대한 자유로운 접근과 이용을 촉진하고자 ‘OAK 포털’을 운영하고 있는데, OAK 포털에서는 각 기관 리포지터리에 등록된 지식정보에 대한 통합검색 서비스를 제공하고 있다. 이처럼 OAK는 오픈 액세스를 기반으로 국가적 차원에서 다양한 지식정보를 발굴하고, 그에 대한 유통과 공동 활용을 촉진하는 새로운 시도라 할 수 있다.

그러나 사업 초기에 다양한 기관의 요구사항을 수렴하여 범용적으로 사용할 수 있는 기관 리포지터리 시스템을 개발하고자 시도하였음에도 불구하고, 현재 OAK-IR 메타데이터에 대한 개선이 시급한 상황이다. 문제의 원인은 크게 두 가지로 분석된다. 하나는 당초 개발된 OAK-IR 메타데이터 자체가 가진 문제이고, 다른 하나는 이들 메타데이터를 운용하는 과정에서 발생하는 문제라 할 수 있다. 가령, OAK-IR 메타데이터는 참여기관의 특수한 상황이나 요구를 사전에 조율하여 이를 시스템 개발에 적극 반영하고자 하였으나 결과적으로 DSpace에서 정의한 메타데이터 요소를 거의 그대로 준용하였다. 이러한 이유로 OAK-IR은 참여기관에서 원하는 메타데이터 요소를 자유롭게 확장하여 사용하도록 허용하고 있는데, 이로 인해 개별 기관에서 자체 정의하여 사용하는 요소가 매우 다양하고, 중복적이며, 상위요소와 하위요소의 논리적 결합이 부적절한 기형적인 모습을 보이고 있다. 뿐만 아니라 이렇게 구축된 데이터를 효율적으로 제어할 수 있는 제도적·기술적 장치가 마련되어 있지 않아 OAK 포털에서 통합검색의 문제를 야기하는 주요 원인이 되고 있다.

이에 본 연구에서는 OAK 참여기관에서 사용하는 메타데이터를 세밀하게 분석하고, 이어 OAK와 유사한 성격의 리포지터리 사례를 조사하여 OAK 메타데이터 수정(안)을 제안하고

자 한다. 연구는 다음과 같은 순서로 진행하였다. 첫째, OAK 참여기관에서 사용하는 메타데이터 요소에 대한 현황조사를 실시하였다. 현황조사는 OAK 참여기관 중 자료수집이 가능했던 17개 기관을 대상으로 이루어졌다. 분석에 필요한 자료는 이들 기관에서 사용하는 메타데이터 명세서를 활용하되, 필요한 경우 리포지터리를 직접 검색하거나 기관 리포지터리 담당자와의 인터뷰를 통해 보완하였다. 둘째, 국내외 주요 기관 리포지터리의 사례를 분석하였다. 사례분석을 위해 국외의 대표적인 기관 리포지터리인 DSpace, EPrints, BEPress, ETD-db와 국내의 dCollection을 분석대상으로 선정하였으며, 이들 리포지터리에서 사용하는 메타데이터 요소와 입력지침, 입력사례 등을 수집하여 분석하였다. 마지막으로, 현황조사와 사례분석의 결과를 토대로 하여 OAK 메타데이터 수정(안)을 제안하였다. OAK 메타데이터 수정(안)은 본 연구팀에서 초안을 도출한 후 OAK 사업에 참여하고 있는 기관 리포지터리 담당자들과 국립중앙도서관의 OAK 사업 담당부서, 그리고 OAK-IR 시스템 개발자들과의 대면 회의 및 이메일 교환을 통해 조율하는 과정을 거쳐 최종 수정(안)을 제안하였다.

II. OAK 참여기관의 메타데이터 요소 분석

OAK 메타데이터는 DSpace 요소를 기본으로 하되 OAK에서 자체적으로 추가한 요소로 구성되어 있다. DC 15개 요소에 ‘provenance’와 ‘administration’ 요소가 추가된 17개의 상위요소와 62개의 한정어로 구성되어 있는데, DSpace와 비교하면 2개의 상위요소(‘creator’, ‘administration’)와 9개의 한정어가 OAK-IR에 추가된 것이라 볼 수 있다. OAK 참여기관에서는 이러한 OAK 메타데이터를 바탕으로 필요에 따라 다양한 요소를 추가하여 사용하고 있다.

OAK 메타데이터 요소와 참여기관에서 자체적으로 확장하여 사용하는 메타데이터 요소를 비교한 결과는 <표 1>과 같다. 참여기관의 메타데이터 요소는 OAK 참여기관 중 17개 기관을 대상으로 분석하였다. 분석은 Dublin Core Metadata Initiative(2004; 2005a ;2005b ;2012)를 기준으로 참여기관에서 사용하는 모든 메타데이터 요소를 매핑한 다음, ‘OAK 기본 요소’와 ‘참여기관에서 확장하여 사용하는 요소’를 구분하여 표시하였다. OAK 기본 요소는 다시 DSpace 요소를 그대로 사용한 경우 <표 1>에 ‘✓’로, OAK에서 자체적으로 추가한 경우 ‘✓(OAK)’로 구분하여 표시하였다. 또한, 해당 요소가 DC 기본요소에 해당하는 경우에는 ‘DC’로, 참여기관에서 자체적으로 추가하여 사용하는 경우에는 해당 기관명을 직접 표시하였다. 각 요소를 사용하는 기관명은 영문 두문자로 기술하되(17개 기관에 대해 A부터 Q까지 총 17개의 영문 두문자 사용), 3개 이상의 기관에서 동일한 의미로 사용하는 요소에 대해서는 기관명을 모두 나열하는 것이 불가능하여 부득이 ‘다수’라고 표시하였다.

4 한국도서관정보학회지(제48권 제1호)

<표 1> OAK 참여기관의 메타데이터 요소 분석

DC 요소	참여기관 사용 요소		OAK 기본	사용 기관	DC 요소	참여기관 사용 요소		OAK 기본	사용 기관	
	요소	한정어				요소	한정어			
dc.title	title		✓	DC	dc.publisher	publisher		✓	DC	
		alternative	✓	DC			alternative			F
		subtitle		A			location		L	
		translated		A	dc.contributor	contributor		✓	DC	
		Title_245_b		B			author		✓	다수
		Title_245_x		B			advisor		✓	다수
		Title_245_n		B			editor		✓	
	Title_245_p		B			creator		✓		
	Title_sub		B		illustrator		✓			
dc.creator	creator		✓(OAK)	DC		other		✓	F, J	
		affiliation		A	dc.contributor	alternativeAuthor		✓(OAK)	다수	
		othername		A			examiner			A
dc.subject	subject		✓	DC			college			L, I
		ddc	✓	DC			department			L, I
		lcc	✓	DC			altAuthorName			M
		classification	✓				googleAuthor			다수
		lsh	✓	DC			id			L
		mesh	✓	DC			Authors_110_a			B
		other	✓	F			Authors_110_b			B
		kdc		D			Authors_710_a			B
		udc		DC			Authors_700_a			B
		nlsh		C		dc.date	date		✓	DC
	naverCode		E				created		✓	DC
	alternative		F		available			✓	DC	
	city		G		issued			✓	DC	
	code		G		copyright			✓	DC	
	evaluation		G		submitted			✓	DC	
	location		G		accessioned			✓	다수	
	map		G		application			✓(OAK)		
	mapcode		G		regist			✓(OAK)		
	name		G		updated					
	province		G		valid				DC	
	Byeolchi		D		modified				DC	
	Subject_650_a		B		accepted				DC	
	Subject_650_x		B		awarded			A		
	Subject_650_z		B		degree			L		
dc.description	description		✓	DC		CREATE_DATE			B	
		abstract	✓	DC		MODIFY_DATE			B	
		tableOfContents	✓	DC		sctpcdate			K	
		statementOf Responsibility	✓	H, I		tcdte			K	
		sponsorship	✓	A, J	dc.type	type		✓	DC	
	provenance	✓	다수			DCMI Type			DC	
	uri	✓	J			DSpace Type			다수	
	claim	✓(OAK)				MARC Type			D	
	version					RIMS Type			K	
dc.format	summary			E	dc.format	format		✓	DC	
	scopustc			K			medium		✓	DC
	wostc			K			extent		✓	DC
	localRemark			A			mimetype		✓	
	alternativeAbstract			F			page			A
	memberprice			F			citation_startPage			B
	price			F			citation_endPage			B
	state			F						

OAK 메타데이터 요소 분석 및 수정(안) 제안에 관한 연구 5

DC 요소	참여기관 사용 요소		OAK 기본	사용 기관	DC 요소	참여기관 사용 요소		OAK 기본	사용 기관	
	요소	한정어				요소	한정어			
dc.identifier	identifier		✓	DC		relation	requires	✓	DC	
		uri	✓	DC			isRequiredBy			DC
		citation	✓				conformsTo			DC
		govdoc	✓	J			ispartofseries	✓		F
		isbn	✓	F, A			isBasedOn	✓		
		ismn	✓				journal			O
		issn	✓	다수			journaltitle			A
		sici	✓				index (학술지명)			L
		other	✓	다수			volume			다수
		applicationNo	✓(OAK)	다수			issue			다수
		registno	✓(OAK)				page	✓(OAK)		다수
		bibliographicCitation		DC			startpage			A
		controlno		I			lastpage			A
		applno		J			Series_Report_Title			B
		patentno		J, N			Series_Report_No			B
		registrationNo		K			index (등재정보)			A
		islink		J			area			G
		pmid		N			classification			G
		scopusid		A, K			code			G
		wosid		A, K			name			G
		rimsid		K		dc.coverage	coverage	spatial	✓	DC
		major		A				temporal	✓	DC
		thesisdegree		A			rights		✓	DC
		holdings		A		dc.rights		uri	✓	DC
		Citation_090_a		B				license	✓	DC
		Citation_090_b		B				holder		
		Citation_090_c		B				isfree		I
		orcid		N			provenance		✓	DC
		contributor		A			administration	regnum	✓(OAK)	B
		pid		K				RECORD_STATE		B
	uid		K			CLASS_NAME		B		
	slug					FILE_CNT		B		
	country		K			call_no_prefix		B		
	kdc		D			Thumb_url		B		
dc.source	source		✓	DC	extension	naverlinktype		E		
	uri		✓	DC		navercpid		E		
dc.language	language		✓	DC		navertype		E		
	iso639-2		✓	DC		statementOf Responsibility		E		
	rfc3066			DC	dc.coverage					
	Alternative			F		citation	date	A		
dc.relation	relation		✓	DC	해당없음		genre		A	
		uri	✓	DC			identifier			A
		ispartof	✓	DC			issueno			A
		haspart	✓	DC			pages			A
		isformatof	✓	DC			beginingPages			Q
		hasformat				DC		title		A
		isversionof	✓	DC			volume			A
		hasversion	✓	DC			confName			A
		replaces	✓	DC			reviewer			O
		isreplacedby	✓	DC			sector	campus		P
		references				DC		daehak		P
		isreferencedby	✓	DC				department		P
				DC	author	department		O		
				DC		google		A, O		

<표 1>에 보듯이, OAK 참여기관에서 사용하는 메타데이터 요소에는 OAK에서 기본적으로 제공하는 요소 뿐 아니라 자체적으로 확장하여 사용하는 요소가 상당히 많이 포함되어 있다. 그러나 기관에 따라 동일한 의미에 대해 상이한 요소명을 사용하거나¹⁾, 동일한 한정어가 상이한 요소 아래에 확장되어 있는 경우²⁾도 있고, 한정어의 사용이 부적절한 경우도 더러 발견된다. OAK 메타데이터 수정(안) 도출을 위해 현재 기관 리포지터리에서 사용하는 메타데이터 요소를 분석한 결과는 다음과 같다.

1. title

OAK에서는 title 및 title.alternative를 표준 메타데이터 요소로 정의하고 있으나, 참여기관에서는 title 요소 아래 subtitle(하위표제), translated(번역표제), title_245_b(표제관련 정보), title_245_x(대등표제), title_245_n(권차), title_245_p(권차표제) 등과 같은 다양한 한정어로 표제를 구분하고 있다. 표제 정보 가운데 대체표제와 원표제는 데이터의 성격이 상이하고, 관계정보의 처리를 위해 구분하여 기술할 필요가 있고, title에 혼재되어 있는 권차(편차), 권제(편제)도 별도의 한정어를 추가하여 구분 입력하여야 할 것으로 판단된다.

2. creator

대부분의 OAK 참여기관에서 creator 대신 contributor에 다양한 한정어를 세분하여 저자 관련 정보를 입력하고 있다. 다만, A기관에서는 유일하게, creator에 저자명을 기술하고, 그 아래 저자의 소속(affiliation)이나 영문명(othertext)을 한정어로 사용하고 있다. 이는 덤다운의 원칙에 위배하는 것이라고 볼 수 있는데 덤다운 원칙은 한정어의 설계에서 한정어를 없애더라도 기록된 값과 데이터 항목 간에 모순이 생겨서는 안 된다는 원칙이다(谷口祥一, 綠川信之 2007, 270). 따라서 저자를 기록하는 creator에 해당 저자를 직접 지시하는 것이라 보기 어려운 '소속기관'이나 '주소' 등의 한정어를 사용하는 것은 적절하지 않다고 볼 수 있다.

3. subject

OAK에서는 주제 정보 기술을 위한 전거스킴으로 DDC, LCC, UDC, LCSH, MESH를 한정어로 정의하고, 그 외 분류체계나 주제어의 수용을 위해 기타 분류체계를 포괄하는 'classification'과 기타 주제명표에 포괄하는 'other'를 한정어에 추가하였다. 반면, 참여기관

- 1) 예: 부차적 표제에 대해 A기관에서는 title.subtitle을, B기관에서는 title.Title_245_b를 별도로 정의하여 적용하고 있음.
- 2) 예: '구글 저자명'의 기술을 위해 다수의 기관에서는 contributor 요소 아래 한정어 googleAuthor를 신설하여 사용하고 있으나, A, O기관에서는 author 요소를 신설한 후 그 아래 한정어 google을 사용하고 있음.

에서는 각 기관에서 사용하는 다양한 분류체계를 추가하고 있으므로 다양한 주제 관련 한정어를 추가할 수 있도록 확장 방안이 마련되어야 한다.

4. description

OAK에서 일반적인 주기사항은 description에 기술하고, 그 한정어로 abstract(초록·요약), tableOfContents(목차), statementOfResponsibility(책임사항), sponsorship(지원기관), claim(특허청구항), version(동료평가), provenance(자원의 소유권) 등을 사용하였다. statementOfResponsibility는 자원에 표시된 문자나 형식 그대로 저자명을 입력하거나 2인 이상의 공저인 경우 자원에 표시된 순서대로 저자명을 기술하는데 사용된다. provenance에는 자원관리와 관련된 변경내역을 기술하는데, 최근 DC에서는 15개 기본 요소 이외에 provenance를 별도의 요소로 확장하고 하지만, OAK 참여기관 대부분은 현재 description의 한정어로 provenance를 적용하고 있다.

5. publisher

OAK에서는 다른 언어 등 복수의 발행사를 입력하고자 할 때 publisher를 반복 적용하도록 정의하고 있다. 그러나 일부 참여기관에서는 다른 언어로 된 발행사의 표기를 위해 한정어 alternative를 두거나(F기관), 발행사가 위치한 국가명이나 지역명을 기술하기 위해 location을 한정어로 사용하고 있다(L기관).

6. contributor

OAK에서는 author(저자), advisor(지도교수), editor(편집자), illustrator(삽화가), other(기타저자) 등 한정어를 세분하고 있다. 여기에, 참여기관에서는 examiner(심사자), college(학교), department(학과) 등을 추가로 적용하고 있어 저자 유형에 대한 한정어의 정리가 반드시 필요한 상황이다. 또한, 저자명의 다양한 표현을 위해 OAK에서 정의한 alternativeAuthor 외에 altAuthorName, googleAuthor와 같은 유사한 한정어도 더러 발견되고 있다. 이외에도 id라는 저자 식별기호를 사용하는 기관도 있는데, 연구자의 고유 식별번호가 점차 중요시되는 상황을 고려할 때, 향후 OAK에서도 저자 식별기호를 위한 요소의 확장이 필요하다고 볼 수 있다.

7. date

OAK에서는 created(제작일), available(이용가능일), issued(발행일), copyright(저작권일), submitted(제출일), accessioned(등록일), application(특허출원일), regist(특허

8 한국도서관정보학회지(제48권 제1호)

등록일)을 사용하고 있다. date 요소는 다른 요소에 비해 많은 한정어를 사용하고 있으나, 한정어가 복잡하고 각각의 사용을 위한 정의가 명확하지 않다는 문제를 가지고 있다. 참여기관에서는 학위논문의 awarded(학위수여일), degree(학위구분)을 사용하고, scptcdate(scopus 피인용정보를 가져온 날짜), tcdate(wos 피인용정보를 가져온 날짜)를 한정어로 사용하는 경우도 있어 이러한 한정어 사용의 필요성을 보다 면밀히 검토해 볼 필요가 있다.

8. type

OAK 참여기관에서는 DCMI type 뿐만 아니라 DSpace type도 널리 사용하고 있다. 또 일부에서는 MARC type, RIMS type과 같은 한정어를 사용하고 있다. 자원유형은 자원유형별 검색에 널리 사용되므로 통일된 방식으로 자원유형을 입력하거나, 적어도 상이한 자원유형을 상호매핑하여 운용하는 방안이 마련되어야 할 것으로 보인다.

9. format

OAK에서 물리적 용기를 medium에, 자원의 크기를 extent에, mimetype을 한정어로 추가하고 있다. mimetype은 medium과 실질적으로 동일한 의미이므로 한정어의 중복으로 볼 수 있다. 참여기관 중 일부는 format 요소에 page(전체 페이지) 또는 citation_startPage(시작페이지), citation_endPage(끝페이지)와 같은 한정어를 사용하여 해당 자원의 페이지를 기술하는 경우도 있어 이에 대한 조정이 필요하다고 볼 수 있다.

10. identifier

identifier는 자원식별자나 인용정보를 기술하는 요소이다. OAK에서는 uri, govdoc, isbn, ismn, issn, sici, other, applicationNo, registNo 등을 다양한 한정어로 추가하고 있다. 참여기관에서는 한정어로 controlno(관리번호), applno(특허출원번호), patentno(특허등록번호), registrationNo(특허등록번호), pmid(pubmed 논문식별자), scopusid(scopus 논문식별자), wosid(wos논문식별자), rimsid(RIMS논문식별자) 등을 사용하고 있다. 최근 자원을 나타내는 식별자를 다양하게 사용하고 있는 추세이므로, OAK에서는 기존 한정어 이외에 다양한 자원식별자 유형을 기술할 수 있도록 한정어를 확장할 필요가 있다. 또한, identifier 기술과 관련하여 가장 많이 발견되는 오류는 기술대상 자료가 학술논문임에도 학술지(journal)의 issn을 identifier에 기술하고 있는 것이다. identifier에는 기술대상 자료와 직접 관련된 식별자, 즉 기술대상 자료가 논문인 경우 논문의 직접적인 식별자를 기술해야 하며, 해당 논문이 수록된 저널의 식별자는 relation의 상위자료(ispartof)에 기술해야 한다. 이러한 오류

를 줄이기 위해서는 이와 관련된 입력지침을 보다 명확히 제시할 필요가 있다.

OAK에서 자원의 인용정보를 citation에 기술하고 있으나, 인용 스타일에 따라 기술방식이 상이하기 때문에 기술의 일관성을 유지하기가 상당히 어렵다. 이에 참여기관에서는 저널명, 권, 호 등으로 데이터를 구분하여 입력하고 있으나 이를 위한 상위요소와 하위요소의 논리적 결합이 맞지 않는 문제가 발생한다. 예를 들어, 일부 기관에서는 인용정보를 relation 요소에 입력하고 있으며, 또 다른 기관에서는 citation이라는 요소 자체를 별도로 신설한 후 title, volume, date, genre, identifier, issueno, page 등의 한정어를 사용하고 있다. 해당자원의 서지 인용정보는 identifier의 citation에 기술하되, 인용정보의 적절한 기술을 위해 현재 사용되는 다양한 한정어에 대한 조정 작업이 필수적으로 이루어져야 할 것이다.

11. source

OAK에서 source 요소의 사용을 권장하지 않고 있고, 참여기관에서도 현재 source 요소를 사용하지 않고 있다.

12. language

자원의 언어에 해당하는 정보를 입력하기 위해 한정어에 언어코드를 기술한다.

13. relation

relation은 기술대상 자료와 다른 자료와의 관계를 표현하기 위한 것으로써 관계어를 기술하고, 관련 자료의 공식적인 식별체계에 준하는 내용 혹은 uri를 기술한다. DC와 OAK의 관계 한정어를 비교하면, OAK에서 hasformat(다른형태(이후)), references(참고문헌), isRequiredBy(요구자료), conformsTo(관련표준)을 사용하지 않는 대신, isPartOfSeries(관련 총서), isBasedOn(정보원 참조)를 추가하고 있다. 특히, OAK에서는 일반적인 상위-하위자료의 관계와 총서관계를 명확히 구분하기 위해 isPartOfSeries를 추가하였다.

<표 2>의 예시와 같이, 저널에 수록된 논문의 경우 기술대상 자료인 논문과 이 논문이 수록된 저널은 isPartOf 관계를 갖는다. 따라서 한국도서관·정보학회지의 ISSN은 relation 요소에 기술해야 한다. 그러나 저널에 수록된 인용정보(저널명, 권, 호, 페이지 등)는 이 논문 자체에 관한 정보이므로 identifier에 기술해야 한다.

〈표 2〉 저널에 수록된 논문 예시

속성	인코딩체계(URI)	속성값	속성값(URI)
dc:title		FRBR형 검색서비스 개발을 위한 이용자 선호도 분석	
dc:contributor.author		이은주	
dc:subject	http://purl.org/dc/terms/DDC	020	
dc:publisher		한국도서관·정보학회	
dc:type	http://purl.org/dc/terms/DSPACEType	Text	
dc:date.issued	http://purl.org/dc/terms/W3CDTF	2016	
dc:relation.isPartOf			urn:ISSN:1225-0902
dc:identifier.bibliographicCitation		한국도서관·정보학회지 47(2), 259-286. (2016)	

14. coverage

자료의 내용과 관련된 공간적 범위(지역)나 시간적 범위(시대)를 기술한다. OAK에서는 spatial과 temporal을 한정어로 적용하고 있다.

15. rights

OAK에서는 rights의 한정어로 license와 uri를 사용하고 있다. 참여기관에서는 한정어로 isfree(원문 유료여부)를 추가하고, 일부 기관에서는 description의 한정어로 price(가격), memberprice(회원가격), state(비매품유무)를 추가하여 사용하고 있다. 이러한 정보는 accessRights 한정어를 사용하여 통합하는 방안이 고려되어야 할 것이다.

16. 기타 요소

참여기관에서는 DC나 OAK 표준 메타데이터 요소 외에 개별 기관에서 필요한 요소를 추가하여 사용하고 있다. 일부 기관에서는 citation이라는 요소 자체를 추가한 후 다양한 한정어를 사용하여 인용정보의 세부사항을 기술하고 있다. 또 일부 기관에서는 sector라는 요소를 신설한 후 소속 캠퍼스, 단과대학, 학과 등을, 그리고 author 요소를 신설한 후 저자의 소속 기관, 다른 문자나 형식으로 된 이름 등을 기술하고 있다. 이처럼 개별 기관에서 추가한 요소에 대해서는 이미 존재하는 요소나 한정어로의 통합 가능성을 면밀하게 검토하고, OAK 참여기관에서 공통적으로 사용하지 않는 경우에는 해당 기관에서만 확장하여 적용하도록 조정하는 방안이 마련되어야 할 것으로 보인다.

이상과 같이 OAK 기관 리포지터리의 메타데이터 사용 현황을 종합 분석한 결과, 각 기관에서는 OAK 메타데이터 요소를 바탕으로 확장 적용하고 있었으나 공통적으로 다음과 같은

문제점을 안고 있는 것으로 드러났다. 첫째, 참여기관에서 요소나 한정어를 확장할 때 덤다운 원칙을 지키지 못하고 있었다. 둘째, 상위요소와 하위요소의 논리적 결합이 맞지 않는 경우가 있었다. 셋째, 동일한 의미를 가진 데이터에 대해 상이한 요소 또는 상이한 한정어를 적용하고 있는 경우가 적지 않았다. 이러한 문제는 OAK 통합검색 서비스에서 데이터 검색 및 디스플레이 문제를 야기하는 원인이 될 소지가 있으므로 즉각적인 수정이 요구된다고 볼 수 있다.

Ⅲ. 국내외 기관 리포지터리의 메타데이터 사례

OAK 참여기관에서 확장하여 사용하는 요소의 OAK 표준 메타데이터로의 통합 및 향후 확장성을 고려한 새로운 요소의 신설 등에 참고하기 위해, 국내외 대표적인 기관 리포지터리에서 사용하는 메타데이터 요소를 분석하였다. 분석을 위한 사례로는 OAK의 목적과 유사하면서 현재 전 세계에서 가장 많이 활용되고 있는 IR 시스템인 DSpace, EPrints, BEPress, ETD-db와 국내 대표적인 기관 리포지터리인 dCollection을 선정하였다. 2014년 현재 전 세계에 3,045개의 기관 리포지터리 시스템이 구축되어 있다. 이들 리포지터리에서 활용하고 있는 소프트웨어는 DSpace(1,225개), EPrints(468개), BEPress(168개), ETD-db(44개), Fedora(40개), OPUS(26개) 순으로 나타난다. 이를 참조하여 본 연구에서는 국외 사례로 상위 4개 리포지터리를 분석대상으로 선정하였다(Repository 66 Home Page). 이들 기관 리포지터리에서 정의한 메타데이터 요소를 DC를 기준으로 상호 매핑한 결과는 다음 <표 3>과 같다.

<표 3> 주요 기관 리포지터리의 메타데이터 요소 비교

DC		DSpace		EPrints	BEPress	ETD-db		dCollection	
요소	한정어	요소	한정어	요소	요소	요소	한정어	요소	한정어
title		title		title*	title article title**	title		title	
	alternative		alternative	alternative title			alternative		alternative subTitle translated
creator		creator		creator	author	creator		creator	
									otherName affiliation whols
subject		subject		subject*		subject		subject	
	lcsch		lcsch						lcsch
	mesh		mesh						
	ddc		ddc						ddc
	lcc		lcc						lcc
	udc		classification						
				keywords*	keywords				kdc
			other						otherSubject

12 한국도서관정보학회지(제48권 제1호)

DC		DSpace		EPrints	BEPress	ETD-db		dCollection	
요소	한정어	요소	한정어	요소	요소	요소	한정어	요소	한정어
description		description		description*		description		description	
	tableofcontents		tableofcontents		abstract		abstract		tableofcontents
	abstract		abstract						abstract
			provenance	provenance					provenance
			sponsorship	funder					sponsorship
				grant number					
				parent project					
			statementofresponsibility						
			uri				release		descriptionURI
					versions*				
					peer reviewed				indexed
				publication status	publication status*				
				data collection method					
				contact					
				administrative note			note		
				additional information					
				divisions*	disciplines	degree	name		
							discipline		
							grantor		
							level		localRemark
					comments				
publisher		publisher		publisher	publisher	publisher		publisher	
							country		
				original data publisher					
				original data publication URL					
				original data publication date					
contributor		contributor		contributor		contributor		contributor	
			author						
			advisor	corporate reators			role		advisor
			editor						editor
			illustrator						illustrator
									examiner
			other						otherContributor name(학회)
									location
				depositor*					
date		date		collection dates	date/date range	date		date	
	created		created		embargo period**				created
	valid								
	available		available						
	issued		issued						issued
	modified								modified
			submitted		update submission				submitted
			accessioned						accepted
									awarded
			copyright						copyright
				deposit date*					

DC		DSpace		EPrints 요소	BEPress 요소	ETD-db		dCollection	
요소	한정어	요소	한정어			요소	한정어	요소	한정어
type		type		data type	document type	type		type	subType thesisDegree specialty
format		format			file name duration**	format		format	extent medium mimetype page description
	extent medium		extent medium mimetype						
					additional file geolocation**				
identifier		identifier		unique resource identifier*	resource identifier	identifier		identifier	
	bibliographicCitation		citation		original citation recommended citation				citationTitle citationIdentifier citationGenre citationOfIssueNo citationVolume citationNumber citationTissueNo citationPages citationStartPage citationEndPage citationConference Name citationConference Number citationConference Place citationConference Date citationEdition citationAuthor place duration publisher URIType URIdentity ISBN ISSN SICI
					volume issue**				
					first page** last page**				
			uri isbn issn sici ismn govdoc						KESIRidentifier otheridentifier holdings
source		source						source	
			uri						

14 한국도서관·정보학회지(제48권 제1호)

DC		DSpace		EPrints 요소	BEPress 요소	ETD-db		dCollection	
요소	한정어	요소	한정어			요소	한정어	요소	한정어
language		language		resource language		language		language	
			iso						iso
relation		relation		related resources	related resources			relation	
	is version of		isversionof						isVersionOf
	Has Version		hasversion		previous version**				hasVersion
	Is Replaced By		isreplacedby						isReplacedBy
	Replaces		replaces						replaces
	Is Required By								
	Requires		requires						
	Is Part Of		ispartof						isPartOf
	Has Part		haspart						hasPart
	Is Referenced By		isreferencedby						
References								reference	
Is Format Of		isformatof						isFormatOf	
Has Format								hasFormat	
			isbasedon					isBasedOn	
			ispartofseries						
uri		uri							
coverage		coverage				coverage		coverage	
	spatial		spatial	geographic coverage					spatial
	temporal		temporal	temporal extent					temporal
rights		rights		copyright		rights		rights	
			license		license				
			uri	legal and ethical issues	public domain statement**				

* EPrints 메타데이터 요소는 핵심요소와 상세요소로 구분되며, *표시는 핵심요소를 표시한 것임.

** BEPress 메타데이터 요소 중 **표시는 학술논문에만 적용되는 요소임.

DSpace 메타데이터는 DC 15개 요소에 기초하여, 15개의 요소와 53개의 하위요소(한정어)로 구성되어 있다(The DSpace Developer Team 2016). DC의 15개 요소 중 creator는 contributor 요소 아래 author, advisor, editor 등과 같은 한정어를 세분하여 사용하기 때문에 거의 사용되지 않으며, source는 하베스팅되는 자료에 한해서만 사용되고 있다. DSpace는 DC와 큰 차이는 없지만, description, contributor, date 요소에 해당하는 데이터 값을 구분하여 기술하기 위해 한정어를 추가하였고, subject, identifier에 데이터 값의 근원이 되는 통제어휘의 유형을 보다 세분화하였다는 특징을 가지고 있다.

EPrints 메타데이터 요소는 영국 에식스대학교(University of Essex)의 UK data

Archive를 사례로 도출하였다. 메타데이터 요소를 핵심요소에 해당하는 'EPrint Core'와 기타 요소에 해당하는 'EPrint Detail'로 구분하고 있다(Ensom and Wolton 2012, 1-2). 핵심요소로 title(표제), URI(식별자), description(설명), keywords(키워드), subject(주제어), divisions(부서(대학, 학과, 연구소 등)), depositor(데이터 입력기관 ID) 등 7개 요소를 설정하고 있다. EPrints 메타데이터 요소는 DC에 기초하지 않았기 때문에 DC 15개 요소와 DSpace 요소와 상당한 차이를 보인다. 가령, 주제는 통제어휘와 비통제어휘로 구분하여 subject과 keywords 요소에 따로 입력하도록 설정하고 있으며, 대학 연구물의 특성을 반영하여 funder(연구비 지원기관), grant number(과제번호), parent project(연구프로젝트명) 등 연구비 관련 요소가 상세하게 마련되어 있다. 또한, publisher(발행자) 외에 original data publisher(원문의 발행자), publication status(발행 구분), original data publication date(원문의 발행일자) 등의 요소도 가지고 있다.

BEPress 메타데이터는 DC 요소 및 한정어를 근간으로 하여 기본 메타데이터 요소를 설정하였는데, 이러한 기본 메타데이터 외에 자원유형(단행본, 저널, 학술논문, 이미지 등)별로 추가된 요소도 있다(Otterbein University 2015). 가령, 학술논문에는 단행본이나 이미지에 없는 pre-print, post-print, published와 같은 'versions', 출판 여부를 표시하는 'publication status', 게재 유예기간을 나타내는 'embargo period', 내용상의 변화가 있는 경우 원본 버전을 기술하는 'previous version' 등이 추가되어 있다. 다른 기관 리포지터리의 메타데이터와 비교하여 BEPress에만 존재하는 요소로는 동료평가 여부를 판단하는 'peer reviewed', 저작권 유효기간이 지났거나 포기한 저작물을 표시하는 'public domain statement', 관련 웹사이트나 관련 의견 등을 안내하는 'comments' 등이 있다. 또한, 인용정보에 해당하는 요소가 본래 출판된 자원의 인용정보를 기술하는 'original citation'과 다른 요소에 기술된 내용을 토대로 인용정보를 자동 생성하는 'recommended citation'으로 구분되어 있다. 이에 비해, identifier와 relation 요소에 다양한 한정어가 개발되어 있는 DC나 DSpace와 달리, BEPress는 한정어나 기타 관련 요소 없이 'resource identifier'와 'related resources'만 사용한다는 점이 특이하다.

ETD-db의 메타데이터 요소는 DC를 바탕으로 개발되었기 때문에 DC와 거의 유사하다. 다만, description의 한정어로 note(주기)와 release(자원의 버전), publisher의 한정어로 country(발행사가 위치한 국가), contributor의 한정어로 role(역할어)가 추가되어 있다. 또한, ETD-db가 주로 학위논문을 대상으로 하기 때문에 학위논문과 관련된 요소인 'degree'가 신설되어 있고, degree 요소에 name(학위이름), level(학위구분), discipline(학문분야), grantor(학위수여기관)과 같은 다양한 한정어가 적용되고 있다는 점도 ETD-db만이 가지고 있는 특성이다.

KERIS에서 개발한 dCollection 메타데이터 SOMS는 DSpace와 가장 유사하면서도

DSpace에 적용된 한정어를 수정·변경하여 적용하였다(한국교육학술정보원 2005; 정연경, 이나니, 이미화 2007). title에서 대체표제(alternative) 이외에 다양한 표제를 기술하도록 subtitle(부표제)과 translated(번역표제)를 한정어에 추가하였고, creator에도 otherName(다른 언어로 된 저자명), affiliation(저자의 소속기관), WhoIs(저자 이메일)와 같은 한정어를 추가하였다. subject 요소와 관련하여서는 국내에서 거의 사용하지 않는 MeSH와 UDC를 한정어에서 제외하는 대신 KDC와 otherSubject를 추가하였다. description 요소의 한정어는 DSpace와 유사하게 정의하되, 학술지 등재여부를 기술하기 위한 'indexed'를 추가하였다. 또한, 학위논문 정보를 세분화하기 위해 type 요소의 한정어로 thesisDegree(학위명)과 specialty(세부전공)을 확장 적용하였다. SOMS의 가장 두드러진 특징은 identifier 요소에서 나타나는데, 인용정보를 기술하기 위해 상당히 많은 한정어를 적용하는 한편, uri, isbn, issn, sici, KERISidentifier, otherIdentifier와 같은 다양한 식별자 체계를 한정어로 사용하고 있다는 점이다.

지금까지 살펴본 5개 기관 리포지터리의 사례에서는 각 사례별 특징이나 구축 목적에 따라 다양한 메타데이터 요소를 사용하고 있음을 볼 수 있었다. DC 15개 요소와 비교해 보면, DSpace가 DC를 가장 잘 반영하면서도 대상자원의 특성을 구분하여 기술하기 위해 다양한 한정어를 사용하고 있었다. 또한, 기관 리포지터리의 구축 목적에 맞춰 EPrints는 연구과제나 프로젝트 정보와 관련된 요소를, ETD-db는 학위논문과 관련된 요소를 중점적으로 확장 적용하고 있다는 사실도 발견할 수 있었다. 메타데이터 요소의 적용방식과 관련하여서는 EPrints가 핵심요소와 상세요소로 구분하는 반면, BEPress는 기본요소를 공통적으로 적용하면서 자원유형별로 특화된 요소를 추가하여 적용하고 있었다. 국내 dCollection에 적용되는 SOMS는 DSpace를 대폭 참조하였지만 국내 실정에 맞춰 요소 및 한정어를 변경하거나 확장 적용하고 있음도 뚜렷하게 드러났다. 이러한 사례를 참조하여, OAK 메타데이터도 참여기관의 특성이나 요구에 맞춰 참여기관에서 중요하게 사용되는 요소나 한정어는 수용하되, 유사한 요소나 한정어는 통합하고, 특히 부적절한 한정어의 사용은 대대적으로 수정해야 할 것으로 보인다.

IV. OAK 메타데이터 수정(안) 제안

1. OAK 메타데이터 수정(안)의 도출 근거

OAK 참여기관의 메타데이터 사용 현황과 국내외 기관 리포지터리의 사례조사를 바탕으로 하여, 본 연구팀에서는 기존 메타데이터 요소의 변경, 통합, 신설, 삭제 등을 포함하는 OAK 메타데이터 수정 초안을 작성하였다. 이를 토대로 OAK 사업 총괄 부서와 시스템 개발자, 그리고 개별 기관 리포지터리의 담당자들을 대상으로 대면 회의 및 이메일로 초안에 대한 세밀

한 검토 작업을 실시한 후 최종적인 수정(안) 및 확장 적용방안을 모색하였다. 수정(안)의 주요 내용을 제시하면 다음과 같다.

먼저, 기존 OAK 메타데이터는 DC 및 DSpace에 기초하여 요소 및 각 요소별 한정어를 정의하고 있다. 여기에, 참여기관 자체적으로 추가한 요소 및 한정어도 상당수 존재한다. 이 가운데, OAK 및 참여기관에서 정의한 한정어를 살펴보면, 각 요소의 의미를 세분하거나, 요소 값을 기술하는데 사용되는 통제어의 유형(encoding scheme)을 지시하기 위한 목적이 혼용되어 있다. 따라서 수정(안)에서는 기존의 한정어를 ‘요소구분’과 ‘적용스킴’으로 구분한 후 요소구분에는 그 요소에 해당하는 데이터의 값을 구분하여 기술하는데, 그리고 적용스킴에는 요소 값과 관련된 데이터의 유형이나 특성을 표시하는데 사용되도록 의미를 분리하였다. 예를 들어, 표제 정보는 title.alternative(대체표제), title.original(원표제), title.partName(권차표제) 등으로 구분하는 ‘요소구분’이 필요하고, 주제어는 주제어 값을 표현하기 위한 DDC, KDC, LCSH, NLSH 등과 같은 ‘적용스킴’이 필요하다. 이러한 방식은 현재와 같이 참여기관에서 다양한 한정어를 무리하게 확장할 필요 없이, OAK-IR 시스템에 원하는 통제어의 유형만 간단히 추가한 후 해당 데이터를 입력할 수 있다는 점에서 시스템의 확장성과 사용성을 보장하는데 보다 효과적인 방안이라 할 수 있다. 이에 따라 OAK 수정(안)에서는 subjectType(주제어휘), degreeType(학위구분), eprintType(논문의 버전), contributorType(자원생산자 유형), nameIdentifier(이름 식별자 유형) 등과 같은 통제어 리스트를 개발하여 적용하였다.

다음으로, 참여기관에서 추가하여 사용하는 요소 중 의미가 중복되는 경우에는 요소명을 통합하여 수정(안)에 추가하고(예: abstract, summary → abstract), 요소명의 식별력이 떨어지는 경우에는 보다 적절한 요소명으로 변경하였다(예: claim → patentClaim). 또한, OAK 참여기관의 요구를 수렴하고, 더불어 국내외 기관 리포지터리의 사례를 참조하여 새로운 요소를 추가하고(예: subject.keyword(저자 제공 키워드), description.degree(학위종류), date.valid(유효일자) 등), 통제어의 적용으로 인해 기존의 한정어가 불필요하거나 기존에 부적절하게 사용되었던 한정어는 삭제하였다(예: format.mimetype → 삭제, description.uri → 삭제). 뿐만 아니라 덤다운의 원칙에 명확히 위배된다고 판단되는 요소는 적절하게 이동 배치하였다(예: description.price → rights.accessRights).

또한, 최근 자원과 관련된 식별자를 많이 사용하고 있어 이를 적극 수용하고자 하였다. 가령, identifier에 기존에 한정어에 포함되어 있던 uri, isbn, ismn, issn, sici, govdoc, patentRegistrationNumber, patentApplicationNumber 외에 보다 다양한 자원식별자 유형을 기술할 수 있도록 확장하였다. DataCite(2015)에 따르면, ARK(아카이브자원의 식별자), arXiv(arXiv 리포지터리), EAN13(국제논문번호), EISSN(전자 ISSN), bibcode(천체물리학데이터 시스템 서지코드), PMID, PURL 등과 같은 다양한 식별자가 사용되고 있는데

(DataCite 2015), 이를 참조하여 lissn, doi, istc, uci, url, urn, pmid, scopusid, wosid 등과 같은 식별자를 추가하고, 그 외 로컬식별자를 사용할 수 있도록 localId를 추가하였다.

한편, 기술대상 자료와 관련된 다른 자료를 표현하는데 사용되는 요소가 relation이다. relation 요소는 관계어 및 공식적인 식별체계에 준하는 내용을 기술하여 관련 자료를 연계시키는 용도로 사용된다. 현재 OAK에 정의된 relation 요소의 관계어 한정어는 DC나 DSpace와 약간 차이가 있으며, 한정어 간에 의미가 중복되는 경우도 있다. 이에 OAK 수정(안)을 도출하기 위해 DC, DSpace, DC-Lib Application에 제시된 relation 한정어를 상호 비교하였다. DC, DSpace, DC-Lib Application에서는 관계어의 유형 뿐 아니라 그 의미 해석에 있어서도 다소 차이가 있다. OAK 수정(안)에서는 이처럼 다양한 관계유형을 모두 포괄하되 그 의미를 명확히 정리하는 것이 필요할 것으로 판단되었다. 이에 상위자료<isPartOf>, 포함자료<hasPart>, 다른형태(이전)<isFormatOf>, 다른형태(이후)<hasFormat>, 이전버전 <isVersionOf>, 최근버전<hasVersion>, 선행자료 <replaces>, 후속자료<isReplacedBy>, 참고문헌<references>, 인용한 자료<isReferencedBy>, 필수자료<requires>, 요구자료 <isRequiredBy>, 관련표준<conformsTo>을 OAK 수정(안)에 적용할 관계유형으로 정리하였다. 더불어, 현재 OAK에서 적용하고 있는 관련총서 <isPartOfSeries>는 그대로 적용하되, <isBasedOn>은 참고문헌<references>과 거의 동일한 의미로 사용되고 있으므로 <isBasedOn>과 <references>을 통합하는 것으로 조정하였다.

마지막으로, 참여기관마다 상이하게 적용하는 것으로 나타났던 ‘인용정보’에 대한 수정(안)을 마련하였다. 기존 OAK에서는 identifier 요소의 citation 한정어에 인용정보를 기술하도록 정의하고 있었으나, 다수의 참여기관들은 relation 요소 아래 다양한 한정어를 추가하여 적용하고 있는 것으로 나타났다. 이처럼 참여기관에서 자체적으로 인용정보를 확장 적용하다 보니 한정어의 사용에 있어서도 동일한 의미에 대해 상이한 한정어를 사용하거나(예: journal 과 journaltitle), 다른 의미를 가진 동일한 명칭의 한정어도 더러 발견되고(예: index(학술지 명으로 사용)과 index(등재정보 기술)), 한정어의 세분화 정도에서도 차이가 있었다(예: page와 startPage/lastPage). 이러한 문제에 더해, 가장 큰 문제는 인용정보를 relation 요소에 기술하는 것이 부적절하다는 점이었다. 앞서 언급하였듯이 relation은 기술대상자료와 관련된 다른 자료와의 ‘관계’를 기술하기 위해 사용되는 요소이며, 이러한 이유로 DC나 DSpace, BEPress에서도 인용정보를 identifier 요소에 기술하고 있다.

이에 참여기관의 다양한 의견을 수렴하되 OAK-IR 시스템 운용의 측면을 고려하여 인용 정보에 대한 수정(안)은 다음과 같이 도출되었다: (1) 인용정보를 위한 별도의 요소로 citation을 추가한다. (2) 인용정보 기술을 위해 citation 요소 아래 title, volume, number, date, startPage, endPage, conferenceName, conferenceNumber, conferencePlace, conferenceDate, author, edition, place, publisher로 한정어를 세분한다. (3) 인용정보는

citation 요소에 입력하되 여기에 입력된 정보는 관리용으로만 사용하며, citation 요소에 입력된 정보를 통합하여 identifier.bibliographicCitation을 자동 생성한다.³⁾ (4) identifier.bibliographicCitation의 인용정보는 특정 기술방식(예: MLA, APA 등)에 따라 데이터의 결합 순서와 방법을 따로 정할 수 있도록 한다. 인용정보의 입력 및 생성 예시는 다음 <표 4>와 같다.

<표 4> citation 및 identifier 요소의 인용정보

요소	요소구분	입력 값 예시	비고
identifier	bibliographicCitation	한국정치학회보, 4, 2(2002): 211-236	자동생성
citation	title	한국정치학회보	직접입력
	volume	4	직접입력
	number	2	직접입력
	date	2002	직접입력
	startPage	211	직접입력
	endPage	236	직접입력

2. OAK 메타데이터 수정(안)

앞서 검토한 내용을 바탕으로 최종 도출된 OAK 메타데이터 수정(안)은 다음 <표 5>와 같다. 수정(안)은 현행 OAK 메타데이터 요소와 비교하여 제시하였으며, 주요 수정내용과 필수 및 반복여부를 간단히 표시하였다. 더불어, 해당 요소가 OAK-IR 간 유통 또는 OAK 포털에서의 통합검색에 사용되는 경우에는 ‘유통’으로, 그리고 데이터 입력기관에서만 필요로 하거나 관리 용도로 사용되는 요소에 대해서는 ‘관리’로 표시하였다.

<표 5> OAK 메타데이터 수정(안)

현행		수정(안)			수정 내용	필수	반복	비고	
요소	한정어	요소	요소구분	적용스킴					
				용어	구문				
title		title					✓	유통	
	alternative		alternative					✓	유통
			original			추가		✓	유통
			partNumber			추가		✓	유통
			partName			추가		✓	유통
creator		creator						미사용	
subject		subject		<subjectType> KDC DDC LCC UDC LCSH MESH NLSH local other		통제어 리스트 적용		✓	유통
	ddc								
	lcc								
	classification								

3) 해당 자원이 학술지논문인 경우에는 citation.title(표제), citation.volume(권), citation.number(호), citation.date(수록년도), citation.startPage(시작페이지), citation.endPage(끝페이지) 등에 입력된 값을 결합하여 시스템에서 자동 생성한다. 해당 자원이 단행본의 챗터인 경우에는 citation.title(표제), citation.author(저자), citation.date(수록년도), citation.startPage(시작페이지), Citation.endPage(끝페이지), citation.edition(판) 등에 입력된 값을 결합하여 시스템에서 자동 생성한다.

20 한국도서관정보학회지(제48권 제1호)

요소	현행	수정(안)				수정내용	필수	반복	비고
	한정어	요소	요소구분	적용스킴					
				용어	구문				
	lsh mesh other		keyword			추가(저자생성키워드)	✓	유통	
description		description					✓	유통	
	abstract		abstract				✓	유통	
	tableOfContents		tableOfContents				✓	유통	
	statementOfResponsibility		statementOfResponsibility				✓	유통	
	claim		patentClaim			요소명변경	✓	유통	
	sponsorship		sponsorship				✓	유통	
			degree	<degree Type> doctoral master		추가 통제어 리스트 적용		유통	
			eprintVersion	<eprint Type> preprint postprint published		추가 통제어 리스트 적용		유통	
	provenance		provenance				✓	관리	
	uri		-			삭제			
publisher		publisher					✓	유통	
			location			추가	✓	유통	
contributor		contributor		<contributor Type> author advisor editor translator illustrator examiner department reviewer other		통제어 리스트 적용	✓ (해당 시 필 수)	유통	
	author								
	advisor								
	editor								
	creator								
	illustrator								
	other								
			affiliation			추가	✓	유통	
	alternativeAuthor		alternativeName			요소명변경	✓	유통	
			nameIdentifier	<nameIdentifier Type> orcid viaf isni scopusId researcherId localId other		통제어 리스트 적용	✓	유통	
			affiliatedAuthor			추가	✓	관리	
			approver			추가	✓	관리	
date		date					✓	유통	
	created		created		DCMI Period W3C-DTF ISO8601			유통	
			valid			추가			유통
	available		available						유통
	issued		issued						유통
	updated		modified			추가			유통
			dateAccepted			추가			유통
	accessioned		accessioned						유통
	copyright		dateCopyright						유통
	submitted		dateSubmitted						유통
			awarded						유통
	application		application				추가		유통
	regist		registration						유통

현행		수정(안)				수정내용	필수	반복	비고
요소	한정어	요소	요소구분	적용스킴					
				용어	구분				
type		type		<resourceType> DCMI Type MARC genre DSpace local other		통계어 리스트 적용		✓	유통
format		format			IMT			✓	유통
	medium		medium					✓	유통
	extent		extent					✓	유통
	mimetype		-			삭제			
identifier		identifier		<IdentifierType> isbn issn lissn ismn istc doi uci uri url urn govdoc patentRegistration Number patentApplicationN umber sici pmid scopusid wosid localld		통계어 리스트 적용		✓	유통
	uri								
	govdoc								
	isbn								
	ismn								
	issn								
	sici								
	other								
	applicationNo								
	registNo								
	citation		bibliographicCitation			요소명변경 및 타이틀 생성 방식 변경		✓	유통
source		source							미사용
language		language			iso639-2 rfc3066	언어코드 추가		✓	유통
	iso639-2					삭제			
relation		relation			isbn issn lissn ismn istc doi uci uri url urn govdoc patentRegistrationNu mber patentApplicationNum ber sici pmid scopusid wosid localld			✓	유통
	isPartOf		isPartOf					✓	유통
	hasPart		hasPart					✓	유통
	isFormatOf		isFormatOf					✓	유통
			hasFormat			추가		✓	유통
	isVersionOf		isVersionOf					✓	유통
	hasVersion		hasVersion					✓	유통
	replaces		replaces					✓	유통

22 한국도서관정보학회지(제48권 제1호)

현행		수정(안)			수정내용	필수	반복	비고	
요소	한정어	요소	요소구분	적용스킴					
				용어					구분
	isReplacedBy		isReplacedBy				✓	유통	
			references			추가		✓ 유통	
	isReferencedBy		isReferencedBy					✓ 유통	
	require		require					✓ 유통	
			isRequiredBy			추가		✓ 유통	
			conformsTo			추가		✓ 유통	
	isPartOfSeries		isPartOfSeries					✓ 유통	
	isBasedOn		-			삭제			
	uri		-			삭제			
coverage		coverage						✓ 유통	
	spatial		spatial					✓ 유통	
	temporal		temporal					✓ 유통	
rights		rights		uri				✓ 유통	
			accessRights			추가		✓ 유통	
			rightsHolder			추가		✓ 유통	
	license		license	uri				✓ 유통	
	uri					삭제			
		citation	title			추가		✓ 관리	
			volume			추가		✓ 관리	
			number			추가		✓ 관리	
			date	DCMI Period W3C-DTF ISO8601		추가		✓ 관리	
			startPage			추가		✓ 관리	
			endPage			추가		✓ 관리	
			conferenceName			추가		✓ 관리	
			conferenceNumber			추가		✓ 관리	
			conferencePlace			추가		✓ 관리	
			conferenceDate	DCMI Period W3C-DTF ISO860		추가		✓ 관리	
			author			추가		✓ 관리	
			edition			추가		✓ 관리	
			place			추가		✓ 관리	
			publisher			추가		✓ 관리	

V. 결론

본 연구에서는 OAK 참여기관에서 사용하는 다양한 메타데이터 요소를 분석하고, 국내외 대표적인 기관 리포지터리의 메타데이터 사례를 조사한 다음, 이를 토대로 OAK 메타데이터 수정(안)을 제안하였다. 구체적으로, OAK 참여기관의 메타데이터 현황을 조사하기 위해 17개 기관의 사례를 입수하여 분석하였으며, 참여기관에서 사용하는 메타데이터 요소의 용도가 명확하지 않은 경우 기관 리포지터리 담당자와의 인터뷰를 통해 이를 보완하였다. 이어,

DSpace, EPrints, BEPress, ETD-db, dCollection 등 국내외 주요 기관 리포지터리의 사례를 분석한 다음, 관련 실무자들과의 최종적인 검토를 거쳐 OAK 수정(안)을 도출하였다.

OAK 수정(안)에서는 기존 OAK 메타데이터에 포함되지 않았지만 참여기관에서 공통적으로 널리 사용하고 있는 데이터를 ‘표준’ 메타데이터로 공식화시켰으며, 기관에 따라 상이한 요소명을 사용하는 경우에는 통합하고, 부적절한 요소의 사용에 대해서는 변경 또는 삭제하는 방안을 제안하였다. 뿐만 아니라 기존에 참여기관마다 요소 또는 한정어를 임의로 추가하여 사용하던 방식 대신, 다양한 통제어 리스트를 개발하여 ‘적용스킴’으로 활용하도록 함으로써 참여기관에서 원하는 데이터 값을 보다 유연하게 입력할 수 있도록 하였다. 이러한 방식은 OAK 메타데이터 요소의 변경이나 추가 없이 적용스킴의 확장만으로 참여기관의 개별적인 요구에 보다 유연하게 대처하는 효과가 있을 것으로 기대된다. 이외에도, 국내외 리포지터리를 참고하여 메타데이터 요소 및 요소 값과 관련된 자원식별자 등을 확장하였다.

본 연구에서 제안한 OAK 수정(안)은 OAK-IR에 공통적으로 적용할 ‘표준’ 메타데이터의 성격을 지닌다. 따라서 특정 기관에서만 현재 사용되고 있는 요소(가령, 주제를 매우 세분하여 적용하거나(subject.location, subject.map, subject.province 등) 자원의 피인용 지수와 관련된 요소(description.scopustc(Scopus 피인용횟수), date.scptcdate(Scopus 피인용정보의 기준 날짜) 등)는 ‘표준’ 메타데이터에서 제외하였다. 특정 기관에서만 필요로 하는 요소에 대해서는 OAK-IR 표준에 포함하기 보다는 해당 기관에서 적절히 확장하여 사용하고, 이러한 요소는 리포지터리 간의 데이터 유통에서 제외하는 것이 보다 효율적이라고 판단되었기 때문이다.

본 연구의 결과는 향후 OAK 기관 리포지터리의 데이터 공유 및 상호연계를 촉진하는데, 나아가 통합검색의 문제를 개선하는데 크게 기여할 것으로 기대된다. 본 연구에서 제시한 메타데이터 수정(안)이 확정되어 OAK-IR 메타데이터의 표준화가 이루어지고, 더불어 데이터의 품질이 보다 향상되기를 기대한다.

참고문헌

- 정연경, 이나니, 이미화. 2007. 기관 리포지터리 활성화를 위한 메타데이터의 확장 방안 연구. 『한국문헌정보학회지』, 41(1): 323-344.
- 한국교육학술정보원. 2005. 『dCollection 시스템 관리자 운영지침서』. 서울: 동연구원.
- 谷口祥一, 綠川信之. 2007. 『知識資源』. 최석두, 한상길 옮김. 서울: 한울.
- DataCite. 2015. *DataCite Metadata Schema for the Publication and Citation of Research Data*. <http://schema.labs.datacite.org/meta/kernel-3.1/doc/DataCite-MetadataKernel_v3.1.pdf> [cited 2016. 9. 20].

- Dublin Core Metadata Initiative. 2004. *Library Application Profile*. <<http://dublincore.org/documents/library-application-profile/>> [cited 2016. 9. 10].
- Dublin Core Metadata Initiative. 2005a. *Using Dublin Core : Dublin Core Qualifier*. <<http://dublincore.org/documents/usageguide/qualifiers.shtml>> [cited 2016. 9. 10].
- Dublin Core Metadata Initiative. 2005b. *Using Dublin Core : Element*. <<http://www.dublincore.org/documents/usageguide/elements.shtml>> [cited 2016. 9. 10].
- Dublin Core Metadata Initiative. 2012. *DCMI Metadata Terms*. <<http://dublincore.org/documents/2012/06/14/dcmi-terms/?v=elements>> [cited 2016. 9. 15].
- Ensom, Tom and Alexis Wolton. 2012. *RDE Metadata Profile for EPrints*. <http://www.data-archive.ac.uk/media/375386/rde_eprints_metadataprofile.pdf> [cited 2016. 9. 25].
- Otterbein University. 2015. *Digital Commons@Otterbein*. <http://digitalcommons.otterbein.edu/metadata_elements_master_list.pdf> [cited 2016. 9. 25].
- Repository 66 Home Page. <<http://maps.repository66.org/>> [cited 2016. 9. 5].
- The DSpace Developer Team. 2016. *DSpace 5.x Documentation*. <<https://wiki.duraspace.org/display/DSDOC5x>> [cited 2016. 9. 5].

국한문 참고문헌의 영문 표기

(English translation / Romanization of reference originally written in Korean)

- Chung, Yeon-Kyoung, Na-Nee Lee and Mihaw Lee. 2007. "A Study of an Extension of Metadata for Institutional Repository." *Journal of the Korean Society for Library and Information Science*, 41(1): 323-344.
- KERIS. 2005. *dCollection Manual for System Managers*. Seoul: KERIS.