

성공적인 리포지터리의 운영정책에 관한 연구*

A Study on the Operational Policies and Best Practice of Repository

정영미(Youngmi Jung)**

이상기(Sang-Gi Lee)***

초 록

본 연구는 세계의 실 운용중인 리포지터리 중 성공 사례들을 조사하고 그것의 실태와 운영정책을 분석하여 향후의 성공적인 리포지터리 구축과 운영시 유용한 전략적 토대를 마련하고자 하였다. 본 연구의 목적을 달성하기 위해 리포지터리의 성공 사례는 CSIC에서 제공하는 Ranking Web of World Repositories의 Top 10과 JISC에서 지원하는 Registry of Open Access Repositories을 통해 선정되었다. 성공적인 리포지터리의 운영정책 및 관련 자료는 주로 해당 사이트를 통해 수집하였고 큰 범주에서 주제 리포지터리와 기관 리포지터리로 나누었다. 운영정책은 운영 및 지원정책, 수집정책, 관리정책, 배포 및 이용활성화정책, 장기보존 정책, 기술정책, 그리고 저작권 정책으로 구분하여 분석하였다. 또한 성공적인 주제 리포지터리와 기관 리포지터리의 운영정책의 차이 및 국내와 국외 리포지터리간의 차이점도 살펴보았다.

ABSTRACT

This study investigated and analyzed operating policies and status of the best practices repository in the world to propose the successful strategies of management and establishment of repository. The best practices were selected from the top 10 of the Ranking Web of World Repositories by CSIC and Registry of Open Access Repositories by JISC. These repositories fell into two rough classes, disciplinary and institutional repositories. Policies on operation and support, collection, management, distribution and information services, long-term preservation, technology and copyright were analyzed. Characteristic of disciplinary and institutional repository policies were also identified.

키워드: 리포지터리, 오픈 액세스, 운영정책, 주제 리포지터리, 기관 리포지터리
repository, open access, operational policy, disciplinary repository, institutional repository

* 본 논문은 KISTI 지식리포트(2010.10) "리포지터리 사례분석 및 시사점 도출"의 내용을 일부 수정·보완한 것임.

** 동의대학교 인문대학 문헌정보학과 조교수(yomjung@deu.ac.kr) (책임저자)

*** 한국과학기술정보연구원 정보유통본부 정보서비스실 선임연구원(sklee@kisti.re.kr) (공동저자)

■ 논문접수일자: 2010년 11월 15일 ■ 최초심사일자: 2010년 11월 22일 ■ 게재확정일자: 2010년 11월 30일
■ 정보관리학회지, 27(4): 131-152, 2010. [DOI:10.3743/KOSIM.2010.27.4.131]

1. 서론

전통적으로 출판사는 학술정보의 생산을 담당하고 도서관과 정보서비스 기관들은 이들 정보의 유통을 담당해왔다. 하지만 오늘날 학술정보의 급격한 증가, 인쇄저널과 전자저널 가격상승 및 시장변화에 따른 도서관의 재정적 압박, 그리고 디지털 학술정보의 장기보존에 대한 불확실성 등은 학술정보 생산과 유통 구조에 새로운 변화를 요구하고 있다.

최근 몇 년간 이런 요구와 더불어 새로운 학술정보 유통 패러다임인 오픈액세스에 대한 관심이 고조되고 있다. 이러한 최근의 경향은 전 세계적으로 오픈액세스 기반 리포지터리 구축의 원동력이 되고 있다. 해외의 경우 대학이나 연구소, 학회 그리고 연구자들의 자발적인 구성체 등을 통해 이미 많은 리포지터리들이 구축되었고 이들 중 몇몇은 정부지원과 연구자들의 자발적 참여로 상당량의 학술정보를 확보하여 그 이용이 활성화되고 있다.

하지만 국내 리포지터리의 경우 아직 초기단계로 구축현황도 미비할 뿐만 아니라 전 세계적으로 확산되어 이용되고 있는 성공적인 리포지터리 사례는 거의 없다. 세계 리포지터리의 순위를 제공하는 Ranking Web of World Repositories의 결과를 살펴보면 Top 800 내에 국내에서 구축한 리포지터리는 서울대학교의 S-Space와 KAIST의 KOASAS 두 개만이 500위권에 랭크되어 있을 뿐이다. 또한 국내에서는 리포지터리에 대한 전반적인 인식이 부족하고 정부지원도 미흡한 실정이다. 그나마 최근인 2010년 6월에 리포지터리 확산을 위한 노력의 일환으로 한국과학기술정보연구원(KISTI)에서 개발한

한국형 리포지터리인 OAK-IR(Open Access Korea-Institutional Repository)을 기반으로 기관 리포지터리 시범사업이 진행 중에 있어 그 향방을 기대해볼만하다.

이에 본 연구는 국내 리포지터리 구축 및 구축된 리포지터리를 활성화할 수 있는 운영 전략을 도출하기 위해 해외의 성공적인 리포지터리 사례를 중심으로 이들의 운영정책들을 조사·분석하였다. 또한 리포지터리를 크게 주제 리포지터리와 기관 리포지터리로 나누어 그 차이점을 살펴보고 국내와 국외 성공 사례의 차이점도 분석하였다.

2. 이론적 배경

2.1 리포지터리 개념 및 역할

리포지터리는 오픈액세스를 실현하기 위한 중요한 토대로 일반적으로 대학, 연구소 등에서 생산된 학술정보를 아카이빙하여 보존하고 확산하기 위해 개발한 지식저장소로 알려져 있다. 하지만 리포지터리의 정의는 그것의 범위와 역할규정에 따라 다음과 같이 다양할 수 있다.

Lynch(2003)는 리포지터리를 대학이나 기관에서 구성원들에 의해 생산된 디지털자료의 관리 및 배포를 위하여 이용자들에게 제공하는 일련의 서비스라 정의하고 있으며, 가장 필수적인 것은 이들 디지털 자료들에 대한 조직적인 관리 의무로 편목, 접근, 배포뿐만 아니라 장기적인 보존을 언급하였다.

SPARC 보고서에서 Crow(2002)는 리포지터리를 하나 또는 다수의 대학이나 기관의 지

적생산물을 수집하여 보존하는 디지털 서고로 정의하였다.

또한 위키백과(2010)는 이것을 대학이나 연구기관의 구성원들이 생산한 학술정보를 공공의 시각에서 수집하고, 축적하여 학내외의 이용자들에게 무료로 제공하는 인터넷상의 전자서고라고 정의하고 있으며, 또 다른 의미로 개방형으로 운영되는 디지털 개체(digital objects)로 정의하였다.

이들의 정의는 대부분이 기관 리포지터리(institutional repository)에 초점을 맞추고 있으나, 특정 기관에 제한되지 않고 하나나 둘 이상의 주제 분야의 연구자들이 학회나 연구모임 등을 통해 자발적으로 구축한 주제 리포지터리(disciplinary repository)도 다수 존재한다. 본 연구에서는 이 둘 다의 리포지터리를 모두 포함한다.

이러한 리포지터리는 학술정보 유통의 새로운 패러다임이자 차세대 모델로 출판사나 배급자가 수행했던 역할의 일부를 대체할 수 있을 것으로 기대되고 있다. 또한 리포지터리는 운영기관의 위상과 이를 나타내는 지표 역할을 수행할 수도 있다(Crow 2002). 황혜경(2004)은 리포지터리의 역할에 대해 위에서 언급한 두 가지 외에도 추가로 정보에 대한 주제 게이트웨이 역할을 언급하였다.

Rumsey(2006, 182)의 연구에서는 영국 고등교육기관들의 오픈액세스 기반 리포지터리 운영의 목적을 밝히고 있는데 그것은 연구 성과물의 보존과 디지털 객체의 관리, 다양한 버전의 서로 다른 출판물 유형을 취급할 수 있는 메커니즘, 가시성의 증가, 외부의 요구에 순응하기 위함이다.

위의 거시적인 리포지터리 역할과 목적 외에

도 구축하고 운영하는 주체에 따라 리포지터리의 보다 구체적인 목적과 역할이 정의되기도 한다.

Robinson(2009)은 홍콩교육학원 HKIED(Hong Kong Institute of Education) 리포지터리의 목적을 포괄적으로는 연구기관의 연구를 장려하는 하나의 정책이고, 세부적으로 모든 출판된 연구 성과물에 대한 참조 데이터베이스를 제공하여 연구 역량 강화, 기관내 동료들 간에 서로의 연구 성과를 쉽게 공유하고 그들의 업적을 강조, 기관의 연구 성과물에 대한 가시성 향상, 출판된 저작물에 대해 쉽고 편리한 셀프 아카이빙을 제공하고 온라인상에서 자신의 저작물을 관리, 미 출판된 저작물들의 아카이빙을 위한 플랫폼을 제공, 기관 내 연구 성과물에 쉽게 검색과 접근할 수 있는 융통성 있는 데이터베이스 제공, 그리고 마지막으로 기관 내 구성원들의 집단적인 출판 노력을 보여줄 수 있는 수단을 제공하기 위한 것으로 정의하고 있다.

반면 Lercher(2008)은 루이지애나 주립 대학의 8개 학과 교수들을 대상으로 실시한 서베이를 통해 리포지터리의 가장 비중 있는 역할은 출판되지 않는 것만 가치 있는 자료로서의 접근이라는 결론을 제시하고 있다. 특히 그 자료들에는 회색문헌이나 출판전 기사(preprints)로 데이터 세트, 기술보고서, 회의보고서, 강의노트, 워킹페이퍼, 소프트웨어 초고 등을 포함하였다.

2.2 리포지터리의 평가 기준

세계에는 이미 수많은 리포지터리들이 구축되어 운영되고 있다. 이들 리포지터리들의 운영주체의 대부분이 도서관이 차지하고 있지만 이들은 도서관과 분리해서 정의되고 있다. 리

포지터리의 저자 직접 제출(self-archiving)을 지향하는 운영 방식은 도서관과 근본적으로 다르고 이런 이유로 리포지터리의 평가 기준도 도서관의 그것과 다르다.

Shearer(2002)는 기관 리포지터리의 성공을 활성화 정도로 평가하고 두 가지의 정량적인 요소인 제출 활동과 이용률로 접근하고 있다. Xia와 Sun(2007b, 75-78)은 리포지터리의 평가 기준을 정량적인 측면과 더불어 제출자에 대한 제공 정보, 제출된 문헌 수, 원문의 사용가능성, 저자의 태도, 제출 당 비용, 이용 평가, 그리고 상호운용성 등이라고 제시하고 있다.

보다 대중적이고 광범위한 리포지터리의 평가와 그 기준은 <그림 1>의 스페인의 국가연구위원회(CSIC: Consejo Superior de Investigaciones)에서 운영하는 세계 리포지터리 순위 사이트인 Ranking Web of World Repositories에

서 볼 수 있다. 이 사이트에서 제공하는 리포지터리 순위는 <표 1>과 같은 크기, 가시성, 리치 파일, 구글 스칼라 문헌수를 평가 기준으로 사용한 것이다. 본 논문에서는 분석에 필요한 리포지터리 성공 사례는 일차적으로 이 사이트의 Top 10에 포함되어 있는 것이고 최종적으로 분석에 필요한 정책을 명확하게 제시하고 있는 7개의 리포지터리가 선택되었다. 또한 국내의 성공적인 리포지터리 사례를 위해 이 사이트에서 제공하는 800위 내에 유일하게 포함되어 있는 S-Space(538위)와 KOASAS(539위)를 포함하였다.

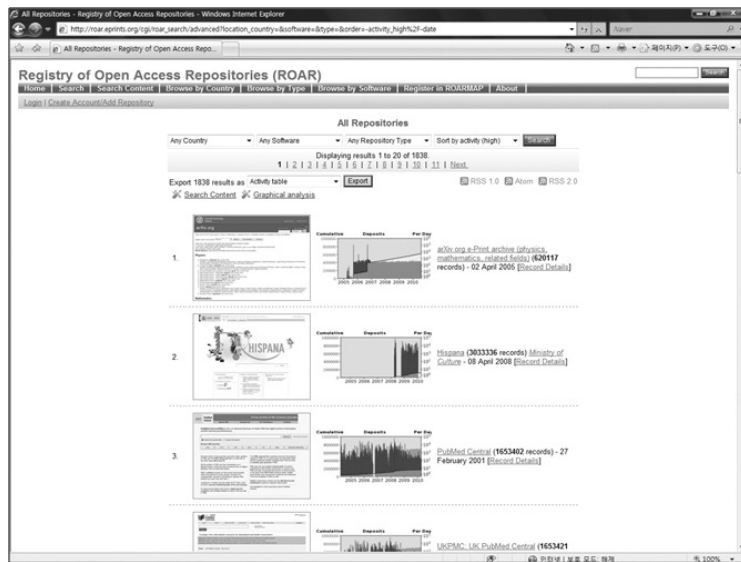
<그림 2>의 JISC가 지원하는 Registry of Open Access Repositories는 또 다른 평가 기준에 의해 세계의 리포지터리들의 순위를 제공하는 사이트이다. 이들은 활동성(이용률)과 보유하고 있는 레코드 건수에 대한 각각의 순위를 제공한다.

WORLD RANK	REPOSITORY	COUNTRY	SIZE	VISIBILITY	POSITION	RICH FILES	SCHOLAR
1	Arxiv.org e-Print Archive		4	2	2	3	
2	CiteSeerX		1	1	411	2	
3	Scientific and Technical Information Network		7	7	1	4	
4	Social Science Research Network		6	3	51	6	
5	Research Papers in Economics		2	6	140	5	
6	Smithsonian/NASA Astrophysics Data System		3	4	539	1	
7	Hal CNRS		18	9	5	8	
8	Institut National de Recherche en Informatique et en Automatique Archive Ouverte		19	11	26	14	
9	CERN Document Server		11	16	40	11	
10	MIT Dspace		14	34	7	9	

<그림 1> Ranking Web of World Repositories [cited 2010.6.10]

〈표 1〉 Ranking Web of World Repositories의 평가 기준

평가 기준	내용	가중치
크기(S)	세계적으로 가장 큰 검색엔진 4개(Google, Yahoo, Live Search, Exalead)에서 발견되는 페이지 수	20%
가시성(V)	Yahoo와 Exalead를 통해 확보할 수 있는 사이트에 의한 외부링크의 총 수	50%
리치 파일(R)	Google과 Yahoo로부터 발췌한 PDF 포맷의 텍스트 파일 수	15%
구글 스칼라 문헌수(Sc)	Google Scholar를 이용하여 2001년에서 2008년 사이에 출판된 최근 논문들과 총 논문수의 정규화된 평균을 계산	15%
전체 순위 계산	$S(20\%) + V(50\%) + R(15\%) + Sc(15\%)$	



〈그림 2〉 Registry of Open Access Repositories [cited 2010.8.10]

앞에서 살펴본 바와 같이 현재 리포지터리의 성공 여부는 학술성, 원문 정보의 제공 등과 같은 정성적인 부분의 기준들도 있지만 대부분이 이용률과 축적 레코드 건수와 같은 정량적인 수치들로 평가되고 있는 것을 알 수 있었다.

2.3 리포지터리의 성공 요인

리포지터리의 활성화, 즉 성공적인 운영 전략을 모색하기 위해 리포지터리의 성공 사례들을

분석하기에 앞서 그것의 성공 요인 및 성공적인 운영 지침에 대해 이론적으로 고찰하고자 한다.

Shearer(2002, 256-257)는 캐나다 연구도서관협회(CARL: Canadian Association of Research Libraries)의 기관 리포지터리에 관한 파일럿 프로젝트에서 리포지터리의 성공에 영향을 미치는 요인으로 제출 활동, 주제 분야, 캠퍼스에서의 옹호 활동, 아카이빙 정책, 저작권 정책, 콘텐츠 유형, 직원의 지원, 질적 통제 정책, 소프트웨어, 이용 등을 제시하고 있다.

CARL 프로젝트의 일환으로 실시된 또 다른 연구에서 Westell(2006, 213)은 프로젝트에 참여한 리포지터리의 성공 요인을 분석하기 위해 이론적 연구들에 근거하여 운영주체의 권한, 모기관과의 방향성 일치여부, 자금 조달 모델, 디지털화 센터와의 관계, 상호운용성, 양적인 측정(제출 활동, 이용률), 홍보 그리고 보존 전략을 성공 요인으로 채택하였다.

Johnson(2002)은 기관 리포지터리의 운영주체와 환경이 다양하지만 기관 리포지터리의 향상을 위해서는 기관적 정의로 기관 리포지터리의 목적과 범위가 명확해야 하고 콘텐츠의 범위는 품질관리를 위해 학술적인 가치가 있는 정보들로 제한하여 구축해야 하고 수집정책 및 제출 기준을 명시해야 할 필요가 있으며 OAI 국제표준을 준수하여 타 시스템과 상호운용성을 지원해야 한다고 제시하고 있다.

국내에서는 황혜경(2009)이 리포지터리 성공에 미치는 요인을 문헌 및 사례연구, 그리고

면담을 통해 도출하였는데, 그것은 크게 기관장의 적극적인 의지를 뜻하는 조직적 요인과 정책적 요인, 그리고 기술적 요인으로 구분된다. 정책적 요인은 핵심콘텐츠 선정, 품질통제 수준의 최소화, 저작권 보호, 적극적 홍보를 통한 옹호 집단 구성, 장기보존 약속, 인센티브, 의무조항을 포함하고 기술적 요인은 제출방식의 편의성, 상호운용성 지원 항목을 포함한다.

CSIC의 Ranking Web of Repositories는 성공적인 리포지터리 운영을 위한 보다 실질적인 열 가지의 지침을 제공한다. 그것은 리포지터리 도메인명의 유지, 많은 콘텐츠 생성 방안 마련, 콘텐츠 포맷의 쉬운 변환, 인터링킹을 통한 콘텐츠 확보, 메인 페이지의 언어, 표준 포맷, 검색엔진의 친화적 설계, 인기도와 통계 관리, 보존과 지속성 정책, 그리고 표준 메타데이터 제공이다.

앞의 선행연구들에서 밝히고 있는 리포지터리의 성공 요소는 정리해보면 <표 2>와 같다. 그래

<표 2> 리포지터리의 성공 요인

구분	Shearer	Westell	Johnson	황혜경	CSIC
운영 및 지원정책	<ul style="list-style-type: none"> • 캠퍼스에서의 옹호활동 • 직원의 지원 	<ul style="list-style-type: none"> • 운영주체의 권한 • 모기관과의 방향성 일치 • 자금조달 모델 • 디지털화센터와의 관계 	<ul style="list-style-type: none"> • 리포지터리의 목적과 범위 명확 	<ul style="list-style-type: none"> • 기관장의 의지 • 옹호집단 구성 	<ul style="list-style-type: none"> • 도메인명 유지
수집정책	<ul style="list-style-type: none"> • 주제 분야 • 콘텐츠 유형 • 제출활동 	<ul style="list-style-type: none"> • 제출활동 	<ul style="list-style-type: none"> • 콘텐츠의 범위 	<ul style="list-style-type: none"> • 핵심콘텐츠 선정 • 제출방식의 편의성 	<ul style="list-style-type: none"> • 콘텐츠 생성 • 인터링킹
관리정책	<ul style="list-style-type: none"> • 질적 통제 	-	-	<ul style="list-style-type: none"> • 품질통제 수준의 최소화 	<ul style="list-style-type: none"> • 콘텐츠 변환 • 표준 포맷
배포 및 이용활성화정책	<ul style="list-style-type: none"> • 이용활성화 	<ul style="list-style-type: none"> • 홍보 • 이용률 	-	<ul style="list-style-type: none"> • 인센티브 • 의무조항 	<ul style="list-style-type: none"> • 홈페이지 언어 • 검색엔진의 친화적 설계 • 인기와 이용통계 관리 • 표준메타데이터 제공
보존정책	<ul style="list-style-type: none"> • 장기보존 정책 	<ul style="list-style-type: none"> • 보존 전략 	-	<ul style="list-style-type: none"> • 장기보존 약속 	<ul style="list-style-type: none"> • 보존과 지속성 정책
기술정책	<ul style="list-style-type: none"> • 소프트웨어 	<ul style="list-style-type: none"> • 상호운용성 	<ul style="list-style-type: none"> • 상호운용성 	<ul style="list-style-type: none"> • 상호운용성 	-
저작권정책	<ul style="list-style-type: none"> • 저작권 정책 	-	-	<ul style="list-style-type: none"> • 저작권 보호 	-

서 본 연구에서는 리포지터리의 성공 요인을 정책적인 관점에서 운영 및 지원, 수집, 관리, 배포 및 이용활성화, 보존, 기술, 저작권정책의 일곱 가지의 영역으로 나누어 살펴보고자 한다.

3. 리포지터리 성공 사례

3.1 주제 리포지터리

3.1.1 arXiv.org

arXiv는 1991년 8월부터 학술지 발표 이전의 연구논문들을 공유하기 시작한 전자서고이자 배포서버로, 현재 코넬대학도서관(Cornell University Library)에서 운영하고 있는 학회 중심의 범 기관적인 리포지터리이다. 주제 범위는 물리학, 수학, 컴퓨터 공학, 비선형과학, 계량 생물학, 통계학으로 연구자들이 직접 자신의 논문을 arXiv에 등록하고 제출된 논문은 거절하거나 재분류하지 않는 것을 원칙으로 하고 있다. 장기적인 보존을 위해 제출된 논문은 삭제하지 않는다.

또한 arXiv는 장기보존을 위해 업로드 파일 유형을 제한하고 있으며, 지원하는 텍스트 파일 포맷은(La)Tex, AMS(La)Tex, PDFLaTeX, DOCX(Word 2007), PDF, PostScript, HTML with JPEG/PNG/GIF이며, 그림은 PostScript, JPEG, GIF, PNG, PDF이다.

제출되는 논문에 대한 최소한의 품질관리를 위해 제출자 등록과정에서 학회 회원임을 인증하는 절차를 통해 arXiv의 커뮤니티로부터 도움을 받지만, 중개자나 관리자가 제출물의 적절성 여부를 결정할 수 없다. 논문을 제출하기

위해서는 간단한 과정의 로그인 절차가 필요하나 이용에는 이런 과정이 필요 없다.

등록한 이용자에 한해 다양한 포맷(제한 있음)으로 학술논문을 제출할 수 있으며, 제출 절차는 ① 메타데이터(타이틀, 저자, 초록 등) 준비, ② 제출물 심의, ③ 제출물 발표, ④ 추가정보, ⑤ 업데이트이다.

리포지터리를 구축하기 위한 시스템은 자체 개발 시스템을 사용하고 있다.

현재 축적하고 있는 레코드는 636,407건(2010.6.30일자 기준)으로 관심 있는 주제 분야의 신규 등록된 논문에 대해서는 RSS와 이메일 알리미 서비스를 제공한다.

3.1.2 CiteSeerX

CiteSeerX는 자동화된 인용색인 작성 및 이를 통해 인용문헌을 링크하는 디지털도서관이다. 1997년 NEC Research Institute의 Steve Lawrence, Lee Giles와 Kurt Bollacker에 의해 개발되어 2003년부터 펜실베이니아대학의 IST(College of Information Sciences and Technology)에서 서비스하고 있다.

CiteSeerX는 과학과 학술적인 지식접근에 대한 기능성, 유용성, 이용가능성, 비용, 편의성, 유효성, 그리고 적시성 제고 및 과학문헌의 배포향상을 목적으로 구축되었고 주로 컴퓨터와 정보공학에 관한 자료를 중심으로 축적하고 있다. 또한 CiteSeerX는 PostScript와 PDF로 제작된 논문을 자동 색인하기 위한 새로운 방법과 알고리즘을 개발해서 사용하고 있다. 현재 약 1,614,142건의 논문기사와 이들과 링크된 31,317,127건의 인용정보를 서비스하고 있다(2010.6.9일자 기준).

CiteSeerX에서는 자동인용색인뿐만 아니라 인용통계, 참조링킹, 인용상황 서비스, 주지서비스 및 트래킹(tracking), 관련 문헌 제공, 전문 색인, 질의기반 요약, 강력한 검색기능, 논문 하비스팅, 메타데이터 자동추출, 개인서고, RSS, 소셜북마킹, 소셜네트워크, 개인별 검색환경 설정 등의 개인화 서비스를 제공한다.

문서의 수집은 연구자들의 직접 제출과 'cite-seerxbot'이라 불리는 크롤러가 담당한다. 연구자들은 CiteSeerX에 부합하는 주제 콘텐츠라면 무엇이든 제출 가능하며, 제출 전에 동료로부터 사전 검토를 권장하고 있다. 제출을 위해 지원하는 파일포맷은 PDF, PostScript, 압축 파일(zip, gz, z)로 다양하지만 PDF 파일을 권장하고 있다. 모든 콘텐츠는 웹을 통해 일반 이용자가 무료로 접근가능하다.

CiteSeerX는 개인정보 보호정책이 있으며, 검색을 위해서는 별도의 이용자 등록이 필요 없지만, CiteSeerX로부터 정보를 받거나 보낼 때, 특히 질문하거나 의견을 제시하고 정보갱신을 요구할 때는 간단한 등록절차가 필요하다. CiteSeerX는 좀 더 높은 수준의 서비스를 위해 이용자의 관심분야를 저장하는 Cookies 서비스를 제공한다. 본 서비스는 이용자들의 접근패턴을 트래킹하여 사이트 성능을 최적화하기 위한 것이다.

리포지터리를 구축하기 위한 시스템으로는 자체개발 시스템을 사용하고 있다.

3.1.3 Social Science Research Network

SSRN(Social Science Research Network)은 사회과학분야 연구 성과물을 웹상에서 빠르게 배포하기 위한 연구네트워크로 SSEP(Social

Science Electronic Publishing Inc.)에서 운영하고 있다. 현재 약 140개 국가의 400,000명의 연구자들이 참여하여 연구논문을 제공하고 있고 그 외에도 수백 개 저널, 출판사, 그리고 출판사와 협력하는 기관들이 참여하고 있다.

SSRN eLibrary는 두 개의 데이터베이스로 구성되어 있는데, Abstract Database는 288,700개 이상의 학술논문과 출판예정 논문의 초록을 수록하고 있고 Electronic Paper Collection은 239,500개 이상의 다운로드 가능한 PDF 파일을 수록하고 있다(2010.6.14일자 기준). 저자가 찬성하면 논문의 원문을 SSRN eLibrary에 자동으로 포함시켜 전자적인 배포를 지원한다.

SSRN은 자신의 초록과 논문을 직접 제출하고, 자신의 전자저널에 대한 구독자를 관리할 수도 있다. 저자 본인의 e-mail을 입력하면 인증 확인 후 등록 처리되고 PDF 논문은 초록과 함께 저널 인용정보, 제출자와 저작권자가 다른 경우 저작권자로부터 받은 승인서, 저자의 e-mail, 주소, 전화번호, Fax를 함께 제출해야 한다.

SSRN의 eLibrary에 논문이나 초록을 제출하는 것은 무료다. 따라서 자신들의 오래된 논문을 포함하여 모든 논문을 제출하는 것이 가능하고 논문 제출에 어떤 제한도 없다. 논문의 이용은 누구든지 가능하다. 저자가 요청하면 SSRN eLibrary로부터 논문을 즉시 삭제할 수 있고 또한 본인이 제출한 논문을 수정하고자 한다면 로그인을 통해 언제든지 수정할 수 있다. 제출된 논문들은 모두 유일한 SSRN의 URL이 부여된다. 하지만 SSRN의 운영주체는 웹사이트, 정책, 이용 조건의 변경에 대한 최소한의 권리만을 지니고 있다. 그래서 운영주체는 웹사이트에 저장된 어떤 자료도 이동하거나 수정할 수 있는

권한이 없고 이용통계를 조정할 수 없다.

또한 보다 강력한 저작권 보호 정책으로 저자는 SSRN에 자료를 제공하더라도 저작권 권리는 완전하게 유지할 수 있다. 일부 요금이 있는 출판사가 제출한 논문도 존재하지만 기본적으로 SSRN 웹사이트의 서비스와 정보는 상업적인 목적이 아닌 오직 개인적인 용도로만 사용할 수 있다.

이용자의 편의를 위해 쿠키서비스와 RSS 등이 제공된다. 리포지터리 구축시 사용 시스템은 자체개발 시스템이다.

3.1.4 Smithsonian/NASA Astrophysics Data System

ADS(SAO/NASA Astrophysics Data System)는 NASA의 동의하에 SAO(Smithsonian Astrophysical Observatory)에 의해 운영되는 천문학과 물리학 연구자들을 위한 디지털도서관이다. ADS는 현재 약 8,400,000건 이상의 레코드를 보유한 세 개의 데이터베이스를 운영 관리하고 있으며, 그것은 천문학과 천체물리학, 물리학, 그리고 arXiv e-prints가 대상이 된다.

ADS는 전자논문, 데이터 목록, 기록물을 포함하여 풍부한 외부자원으로 링크를 제공하고 있으며, ADS에 포함되어 있는 모든 초록과 논문은 출판사에 의해 저작권이 양도되었기 때문에 이용자들은 무료로 사용 가능하다.

ADS는 데이터베이스의 다양성과 완전성을 위해 NASA의 STI(Scientific and Technical Information group), 저널출판사인 SIMBAD, NED 등의 다양한 소스로부터 저널, 출판물, 심포지엄, 회의록, 그리고 내부적인 보고서 초록을 포함한 각종 데이터를 수집하고 있다.

데이터베이스의 품질 제고를 위해 출판 여부에 따라 데이터베이스 완전한 포함 여부를 결정한다. 제출된 논문 중에서 출판전 기사는 고유한 식별자를 부여하여 인용하고 출판중인 논문은 대체로 심사나 검토과정을 마친 상태이기 때문에 출판된 것과 유사하다고 간주하여 데이터베이스에 포함한다. 데이터베이스 업데이트는 일주일에 한번 이루어진다. 모든 미러 사이트도 동시에 갱신된다.

이용자들이 개인적인 용도로 논문을 이용하거나 다운로드하는 것은 무료이다. 검색의 편의를 위해 다양한 검색 방법을 제공하는데 Basic Query Form, Abstract Query Form, ADS Topic Search, Searching by Reference Information, Current/Unread Journals, Searching the ADS by Email 등이 있다. 리포지터리를 구축하기 위한 시스템은 자체개발 시스템을 사용하고 있다.

3.1.5 Research Papers in Economics

RePEc(Research Papers in Economics)는 경제학 분야 지식정보 유통을 촉진하기 위한 70여 개국 수백 명 연구자들의 협력체이다. RePEc는 경제학 분야의 연구보고서, 논문 그리고 소프트웨어 등의 분산데이터베이스로 RePEc의 모든 자료는 무료로 이용 가능하다.

연구자는 소속되어 있는 부서나 기관의 저장소를 통해 RePEc로 자료를 제출할 수 있으며, 등록되지 않은 기관의 연구자는 MPRA(Munich Personal RePEc Archive)에 제출하면 자동으로 RePEc에 전달된다. 연구 성과물을 제출하기 위해서는 RePEc Author Service를 통해 저자 등록이 필요하며 이것은 제출한 연구 성과

물에 대한 품질 제고를 위해 저자 링크를 포함하기 때문이다.

MPRA를 통한 저자의 직접 제출은 저널이나 책에 수록할 목적으로 저작한 경제학 분야 학술 저작물이면 어떤 언어라도 제출할 수 있다. 또한 저작권 소유자의 동의하에 이전에 출판한 논문도 제출할 수 있다. 제출된 논문은 편집자 검토와 승인을 통해 반입된다.

저자는 제출한 연구 성과물의 다운로드 횟수나 인용정보를 항상 확인할 수 있다. 제출의 활성화를 위해 등록된 저자들에게는 매달 이용현황 및 인용정보를 메일링 서비스하고 있다. 이용자는 특별한 등록절차 없이도 RePEc 자료를 이용할 수 있지만, 논문을 제출하거나 등록된 저자 리스트를 보거나 검색하기 위해서는 반드시 등록 절차가 필요하다.

RePEc은 대학의 연구보고서를 제공하기 위해 미국경제협회(American Economic Association)의 EconLit 데이터베이스와 협력하고 있으며 논문의 원문을 구축하지 않고 분산된 데이터베이스로의 링크를 제공한다. 현재 925,000건이 구축되어 있으며 이중 800,000건 이상이 온라인으로 이용 가능하다. 리포지터리를 구축하기 위한 시스템은 자체개발 시스템을 사용하고 있다. RePEc은 개인의 저장소 구축은 지원하지 않지만 기관의 저장소 구축을 위해서는 관련된 기술을 지원한다.

3.2 기관 리포지터리

3.2.1 DTIC Online

DTIC(Defense Technical Information Center)는 전 세계 리포지터리 중 Google과 Yahoo

를 기준으로 PDF 파일을 가장 많이 보유한 리포지터리로, 지난 65년간 국방 분야 정보서비스의 중심기관으로 역할을 해왔다. DTIC는 미국 국방성(Department of Defense)에서 운영 중이며, 연구개발을 지원하기 위해 과학기술정보를 수집, 저장, 검색, 배포하는 국가 정보센터다. DTIC 조력자로는 미국 국방성 조직, 연계 조직, 대학, 과학기술적 활동에 참여하는 비영리 조직, 미국 정부 에이전시 그리고 외국정부가 포함된다. DTIC는 이들 모두에게 연구 성과물의 제출을 권고하고 있으며, 문서 제출 방식은 오프라인과 온라인 두 가지 방식을 지원한다. 특히 DTIC Online은 저자 직접 제출을 지향하며 MS Word, Excel, Powerpoint, WordPerfect, ASCII 텍스트, Rich 텍스트, PDF 등 다양한 포맷을 지원한다.

DTIC는 회원 등록시 자격규정을 두고 있는데 이것은 DTIC의 웹서비스를 이용하거나 접근할 수 있는 수준을 정하는데 활용한다. DTIC의 제품과 서비스는 미국 국방성 직원, 정부직원, 정부 도급자, 하청 도급자, 잠재적인 도급자 모두 이용할 수 있다.

구축하고 있는 리포지터리의 주제 범위는 생물학, 의학, 환경공학, 행동과학, 그리고 사회과학을 포함한 국방 관련 모든 분야를 포함하고 있으며, 미국 국방성의 지침이나 지시, 예산정보, 회의나 심포지엄 자료, 특허와 특허제품, 그리고 국방 커뮤니티에게 흥미로운 자료 모두를 포함한다. DTIC는 이런 자원들과 서비스를 위해 세 개의 웹사이트를 운영하고 있는데 그것은 DTIC Online, DTIC Online Access Controlled, 그리고 DTIC Online Classified다. 이 사이트들은 이용자의 등록 가능 자격이 상이하고, 포

함하고 있는 정보 범위와 접근도 각기 다르다. DTIC Online은 일반인을 위한 것으로 별도의 등록절차가 불필요하지만 DTIC Online Access Controlled는 이용을 위해서 등록이 필요하고, 미국 국방성내 연구 활동에 관한 정보를 제공하는 것으로 미국 국방성 사이트, 수집자료, 기술 보고서, 예산정보, 저널이나 회의자료, 전자책, 진행 중인 연구과제, 표준과 지침 등을 포함하고 있다. 이용에 있어 가장 제한되어 있는 DTIC Online Classified는 미국방성내외의 관계자들에게만 개방되어 있어 기관망 내에서만 무료로 이용 가능하다. 여기에는 2백만 개 이상의 DTIC 기술보고서와 DTIC Online Access Controlled에서는 제공하지 않는 30만개 이상의 연구요약 전문이 포함되어 있다.

3.2.2 CERN Document Server

CERN Document Server는 국제적인 물리학 분야의 연구소인 CERN의 도서관에서 운영하는 리포지터리로 1991년부터 저자들로부터 직접 제출을 통해 대규모의 리포지터리를 구축해왔다.

CERN이 지향하고 있는 주요 두 가지는 그들의 연구 결과물들에 대한 저자 직접 제출방식의 채택과 오픈 액세스 출판이다. CERN 연구소에서 발생한 물리학 및 관련 분야의 출판전 기사, 논문, 도서, 저널, 사진, 비디오, 그리고 기타의 연구 성과물들이 모두 구축 대상이다. 현재 1,071,816건 레코드가 축적되어 있다(2010. 11.1일자 기준).

리포지터리의 관리주체는 CERN의 도서관이지만 리포지터리에 관한 정책 개발과 운영은 Scientific Policy Board에서 책임을 지고 조력

하고 있다.

원칙적으로 CERN 연구자들은 생산한 연구 성과물들을 도서관이나 그들의 연구 부서에 제출하도록 하고 있으나 그렇지 않은 CERN 소속 연구자들의 연구 성과물은 arXiv.org와 같은 다른 리포지터리를 통해 확보하기도 한다. CERN Document Server의 목표는 소속 연구원들 또는 기관에서 생산된 연구 성과물의 100%를 수집하는 것이다(Proudman 2007, 1). 리포지터리 활동과 관련해서 장서개발 팀의 직원 4명, 한 명의 프로그램 팀장과 서비스 개발을 위한 기술직 4명이 전담 배치되어 있다. 그리고 CERN의 리포지터리 활동은 완전하게 도서관 영역의 한 부분이기 때문에 예산 확보가 안정적이다.

연구자가 연구 성과물을 제출하기 위해서는 반드시 로그인을 한 뒤 문서의 유형에 따라 구분되어 있는 제출 엔트리(Submit Entry)로 이동하여 제출하면 된다. 각 문서의 유형에 따라 구성된 엔트리는 또 다시 하위 카테고리 나뉘어져 있다. 예를 들면 'article'을 선택하면 이것은 다시 'reports'와 'preprints'로 구분하여 제출하게 되어 있다. 연구자가 제출할 카테고리만 선택하면 나머지 부분은 모두 메뉴 방식으로 되어 있어 비교적 제출 절차가 용이하다. 제출된 레코드의 서지 정보는 언제든지 수정 가능하다. 하지만 제출된 파일을 수정하는 것은 몇몇 경우에만 허락된다.

그리고 CERN은 연구 성과물 제출을 고무시키기 위해 이용자를 대상으로 다양한 서비스를 제공하고 있다. 우선적으로 도서관의 목록과 통합되어 있고 리포지터리 구축 소프트웨어인 CDS Invenio를 개발하여 보급함으로써 다른

연구기관의 리포지터리와 상호운용이 가능하다. 부수적으로 XML로 문서 작성이나 변환을 할 수 있는 도구를 제공하고 태깅 서비스를 제공한다.

장기적인 보존의 관점에서 CERN은 모든 파일에 대해 새로운 포맷으로 변환할 준비가 되어 있다.

3.2.3 DSpace@MIT

DSpace@MIT는 MIT(Massachusetts Institute of Technology)교수회에서 대학내의 연구 및 학술적인 결과물을 광범위하게 배포하기 위한 방안으로 구축된 리포지터리이다. MIT 교수진들은 자신들의 학술논문에 대한 권리를 학교에 의무적으로 양도함으로써 대학내외의 많은 이용자들이 MIT 교수들이 저작한 학술 논문을 무료로 사용할 수 있게 하였다.

수집되는 학술 논문의 범위에는 논문, 출판 전 기사, 기술보고서, 토의문서, 전자논문, 각종 연구데이터, 이미지/오디오/비디오, 기타 도서관장서가 포함된다. 초기에는 교수들의 연구 자료만 수집하다가 지금은 연구원들의 저작물 중에서 유용성을 인정받은 것과 MIT 교수가 공저자인 논문도 포함하고 있다. 하지만 품질 제고를 위해 여전히 학생들의 논문은 제외하고 있다.

MIT 소속 교수들은 학술 논문에 대한 저작권을 의무적으로 DSpace@MIT에 이양하기 때문에 만약 출판사의 저작권 문제로 리포지터리에 논문 제출이 허용되지 않는 경우에 대비해서 출판 직전 제출을 시행하고 있다. 또한 그 외에도 출판사와의 저작권 문제에 대한 해결 방안들을 구체적으로 마련하고 있다.

3.2.4 S-Space

S-Space는 국내의 서울대학교 중앙도서관에서 운영하는 기관 리포지터리이다. KAIST의 KOASAS와 더불어 Ranking Web of World Repositories에서 제공하는 리포지터리 800순위 안에 유일하게 포함되어 있는 리포지터리이다. S-Space는 2010년 12월 1일자 기준 538위에 랭크되어 있다.

여기에는 서울대학교 소속 교수, 연구원, 대학원생 및 학부생 누구나 본인들의 학술자료를 제출할 수 있다. 제출대상 자료는 저작권을 보유한 출판전 기사나 출판후 기사, 회의자료, 연구자료 및 프로젝트 보고서, 대학·학과·연구소·학회 발간 저널, 석박사 학위논문, 반증자료 및 미완성 자료, 강의노트, 학부생 연구자료 등이다. 하지만 반드시 제출전에 출판사의 저작권 정책을 확인한 후 오픈 액세스가 허용된 논문만을 제출해야 한다.

연구 성과물은 저자가 직접 제출하거나 또는 S-Space 담당자에게 이메일을 통해 등록을 의뢰할 수 있다. 직접 제출 과정이나 메타데이터 작성 과정이 간단하며 제출 후 관리자 인증을 통해 공개하도록 되어 있다. 최소의 품질 관리를 위해 제출자 신분별 컬렉션을 구분하고 있는데 교수나 연구자는 소속 학과나 연구소의 Community/Collection에 자료를 등록할 수 있고 대학원생 및 학부생은 'Student's Archives'의 'Student's Works'에 자료를 등록할 수 있다.

이용자들은 특별한 로그인 절차없이 자료를 무료로 이용할 수 있으나 저자가 제출시 선택한 공개 범위에 의해 학내공개만을 허용하는 자료들도 있다. 이용자들은 기관 리포지터리의 공식적인 사이트 외에도 Google을 통해서도 검색

색 가능하다.

리포지터리 구축 소프트웨어는 DSpace를 사용하고 있고 메인 페이지뿐만 아니라 하위 세션 페이지까지 영어 버전을 제공하고 있다.

3.2.5 KOASAS

KOASAS(KAIST Open Access Self-Archiving System)는 대학내 과학도서관의 학술 정보개발팀에서 운영하고 있는 기관 리포지터리로 Ranking Web of World Repositories에 2010년 12월 1일자 기준 539위에 랭크되어 있다.

KOASAS는 논문 제출의 신분을 대학 소속 교수들에게만 가능하도록 하고 있고 학생이 위임해서 제출하더라도 논문에 대한 최종 책임자는 교수이어야 한다.

KOASAS는 KAIST 소속 교수들의 디지털화된 연구 성과물 전체를 대상으로 하나 현재는 시범적으로 저널 및 컨퍼런스 게재 논문만을 대상으로 한다. 제출 전에는 반드시 출판사 및 저널별 아카이빙 정책을 확인한 후 오픈 액세스 허용 논문만 제출하도록 하고 있다.

제출된 논문은 관리자의 승인과정을 거쳐 승인된 논문만 정상적으로 보존 이용되고 이용자들이 무료로 이용가능하다. 구축 소프트웨어로는 DSpace를 사용하고 있다.

4. 성공적인 리포지터리의 운영정책 분석

4.1 운영 및 지원정책

본 연구에서 조사 분석한 주제 리포지터리들

은 운영주체와 참여주체로 이원화되어 있는데 이런 경우에 참여주체는 주로 주제별 학회나 커뮤니티들이고 운영주체는 대학내의 도서관이나 연구소로 여기는 사이트 관리 등 최소한의 권리만을 수행하는 것으로 나타났다.

arXiv의 경우, 코넬대학도서관에서 사이트를 관리하고 있지만 실질적인 참여주체는 물리학, 수학, 컴퓨터공학 등 학회중심의 전 세계·범기관적인 연구자 커뮤니티다. 커뮤니티든 코넬대학도서관이든 arXiv에 제출된 논문이나 이용 간섭은 최소화하고 사이트를 지속적으로 유지·관리해야 하는 의무는 강화되어 있다. 코넬대학도서관은 arXiv를 유지하고 지속적인 운영체제를 구축하기 위해 운영자금을 확보하고 새로운 비즈니스 모델을 발굴하기 위해 노력하고 있었다.

특히, SSRN은 운영주체의 권한을 분명하게 제한하고 명문화하고 있다. 웹사이트, 정책, 이용조건 등을 바꿀 수 없고, 제출된 어떤 자료도 운영주체가 삭제, 수정, 이동할 수 없도록 하였다.

RePEc은 연구자가 논문을 중앙에 직접 제출하거나 각 기관에 보급한 RePEc 저장소를 통해 제출할 수 있는 분산 체제다. 따라서 RePEc에서는 각 기관에서 공통적으로 필요한 리포지터리 운영기술 등을 제공하고, 실질적인 운영관리는 개별 기관별로 수행한다.

이와 달리 기관 리포지터리들은 운영주체의 권한이 보다 강화되어 있다. 대표적으로 국방관련 리포지터리인 DTIC는 등록과 이용에 많은 제한이 있으며, CERN은 정책 위원회의 지원하에 소속 연구자들의 연구 성과물을 100% 수집하기 위해 제출에 대한 의무규정을 지니고 있으며 마찬가지로 DSpace@MIT 또한 논문

제출을 의무화하기 위해 교수회의 지원을 받는 매우 강력한 권한이 있는 운영주체를 지니고 있다.

하지만 국내의 기관 리포지터리들에는 이런 강력한 논문 제출 의무화 규정이 포함되어 있지 않고 운영주체의 권한에 대한 상세한 설명들이 누락되어 있다.

요약하면, 운영주체가 대학이든 연구소든 관계없이 학회와 같이 자발적인 참여 커뮤니티가 형성될 수 있는 범기관적 주제 리포지터리는 활성화를 위해 운영주체 권한을 최소화하는 것이 필요하지만, 기관 리포지터리는 소속 구성원들의 참여와 활성화를 위해 제출 의무화 등과 같은 보다 강력한 운영정책이 필요하다.

4.2 수집정책

정보수집 방법에는 이용자 직접 제출 방식과 로봇 등을 이용한 자동수집 방식이 있다. arXiv는 이용자들의 직접 제출만으로 운영되는 리포지터리이고, DSpace@MIT와 같은 기관 리포지터리는 소속 대학교수와 연구원들의 직접 제출을 의무화하여 운영된다. 반면 CiteSeerX는 이용자 직접 제출뿐만 아니라 웹상에 존재하는 유용한 학술논문을 로봇으로 자동 수집한다. DTIC도 미 국방성 관련 에이전시들이 구축한 데이터베이스를 일차적으로 수집하고 이용자들의 직접제출을 통해 추가적으로 구축하는 혼합형이다. SSRN은 자료 수집을 위해 이용자의 직접제출뿐만 아니라 출판사를 통한 초록 확보도 병행하고 있다. ADS는 이용자 직접제출, NASA 데이터베이스 자료수집, 협약을 통해 arXiv 자료를 수집하고 CERN의 경우도 유사

하다. RePEc은 기본적으로는 각 기관저장소에 구축된 자료들을 수집하고, 부가적으로 미등록 기관에 소속된 이용자를 위해 중앙에 직접 제출할 수 있게 하고 있다. 국내의 리포지터리인 S-Space와 KOASAS는 저자의 직접 제출과 직접 제출 의사에 통해 관리자의 조력에 의한 수집만을 포함하고 있다. 기관내 소속 연구자들의 자율적인 의사에 전적으로 의지하고 있어 이를 보완할 수 있는 대안의 수집 방법이 존재하지 않는다.

이와 같이 대부분의 리포지터리는 보다 많은 자료를 수집하기 위해 이용자 직접 제출뿐만 아니라 다양한 자료 수집 방법을 병행하는 것으로 나타났다. RePEc과 같이 많은 양의 원문을 구축하는 대신 원문을 이용할 수 있는 링크를 제공하고 관련 기관, 출판사, 저자 연락처 등의 정보를 포함함으로써 이용증가와 가시성을 높이는 사례도 있다.

하지만 기본적으로 모든 리포지터리에서는 저자 직접 제출을 지향한다. 그래서 저자들의 직접 제출을 고무시키기 위해 제출 절차나 방식이 매우 간소화되어 있다. 특히 주제 리포지터리의 경우 논문 제출에 어떤 제한도 없으며, 등록 절차를 거치면 누구나 논문을 제출할 수 있다. 또한 인쇄출판과 같이 복잡한 심사과정도 없다.

4.3 관리정책

리포지터리에 논문을 제출하는 것은 대부분 개방성과 자율성을 목적으로 하기 때문에 일반적으로 인쇄 출판물이 거치는 동료의 심사나 검토과정이 없다. 다만 최소한의 품질관리를 위해

등록한 이용자만 자료를 제출할 수 있으며, 등록 자료에 대해 질의응답할 수 있도록 최소한의 저자 정보를 제공한다. 물론 DSpace@MIT와 KOASAS와 같이 자료를 제출할 수 있는 사람을 학생들을 제외한 교수와 연구원으로 한정하여 콘텐츠 품질을 관리하는 경우도 있다. S-Space는 제출자의 신분에 따라 제출할 수 있는 컬렉션을 분리하고 있는 경우도 있다.

보다 적극적으로 품질을 관리하는 리포지터리는 RePEc로 개인연구자가 직접 제출하는 경우 가이드라인을 명시하고 이에 부합하지 않는 논문은 편집자 판단에 따라 거절할 수 있다. 하지만 깊이 있는 심사보다는 학술적인지, 리포지터리 주제범위에 부합하는지만 판단한다. arXiv도 제출한 논문을 심의하여 기본적인 제출 논문의 포맷에 적절하지 않은 경우 일부 거절하기도 한다. 하지만 대부분의 리포지터리는 활성화를 위해 논문 제출을 적극 권장하고 있으며 어떤 제한 조치도 두지 않고 있다.

다수의 리포지터리는 제출할 수 있는 문서포맷을 제한하거나 자료의 포맷을 변환할 수 있는 도구를 제공하고 있다. 이는 자료 보존뿐만 아니라 이용의 편리성을 제공하기 위함이다. CiteSeerX는 PDF, PostScript, zip, gzip을 허용하지만 가장 권장하는 포맷은 PDF다. 따라서 파일포맷이 맞지 않은 경우 대개 PDF로 변환하여 저장한다. SSRN은 PDF만 허용하고, ADS는 PDF와 PostScript 파일만 가능하다. 대부분의 리포지터리들이 PDF를 가장 선호하는 이유는 논문에 텍스트뿐만 아니라 이미지나 미디어를 포함할 수 있으며, 전 세계 이용자들이 무료 읽기 프로그램을 이용하여 쉽게 이용할 수 있기 때문이다.

사례로 분석된 모든 리포지터리들이 제출된 논문을 관리하고 접근을 개선하기 위해 논문 제출과 동시에 고유식별자를 부여한다. arXiv는 승인된 각 논문에 고유 ID를 부여하며, 보다 편리하게 메타데이터를 작성할 수 있도록 관련 시스템을 제공한다.

주요 리포지터리의 모든 홈페이지와 서브페이지는 영어로 작성하여 서비스한다. 영어가 자국어가 아닌 경우 자국어와 함께 영어 버전을 별도로 제공한다. 심지어 CERN의 경우에는 20여개의 언어로 된 홈페이지를 제공하고 있다. 대부분의 리포지터리에서 제출되는 논문의 언어에는 제한을 두지 않고 있으며, RePEc와 같이 언어제한이 없음을 정책에 명시하고 있는 경우도 있다.

4.4 배포 및 이용활성화정책

이용의 제한은 오직 DTIC의 일부 데이터베이스와 같은 국방과 관련한 중요 데이터베이스만 해당된다. 대부분의 리포지터리는 오픈 액세스 기반 무료이용 및 확산을 목적으로 구축된 시스템이기 때문에 회원등록이나 로그인 없이 모든 자원을 편리하게 무료로 이용할 수 있다. 즉 상업적인 목적이 아니라면 대부분의 리포지터리들이 이용에 어떤 제한도 두지 않는다. 단지 S-Space의 경우, 자료 제출시 제출자의 의향에 따라 공개 여부를 대학내로 제한할 수 있는 옵션을 가지고 있다.

대부분의 리포지터리에서는 이용의 활성화를 위해 최신정보, 뉴스정보, 이용통계, 인용상황 등을 신속하게 제공하기 위해 RSS와 이메일 알리미 서비스를 제공하고 있다. arXiv는 이용

자 프로파일에 등록된 관심 분야와 관련성 높은 새로운 논문이 입수되면 RSS와 알리미 서비스로 신속하게 제공한다. CiteSeerX는 알리미 서비스를 통해 관심 있는 논문의 인용상황을 제공한다. SSRN는 RSS를 이용하여 새로운 논문이나 뉴스 등을 신속하게 전달하고, RePEc은 알리미 서비스를 통해 이용자가 제출한 논문의 이용통계와 인용정보를 제공한다. 이런 서비스는 최신자료 제공뿐만 아니라 자신이 제출한 논문에 대한 지속적인 관심을 유발시켜 리포지터리를 활성화하는데 많은 도움이 된다.

리포지터리와 같이 자발적인 이용자 참여가 운영의 큰 부분을 차지하는 경우, 실시간 나타나는 이용자수 통계는 더 많은 이용으로 이어지고 더 많은 논문의 제출을 유도할 수 있는 요인이 된다. 그래서 대부분의 리포지터리에서 이용통계 정보를 제공하고 있다.

CiteSeerX, RePEc, ADS는 인용정보 분석 서비스를 제공하고 있는데, 인용정보는 리포지터리의 순위를 높이는데 매우 중요한 요소이다. <표 3>에서와 같이 3개 리포지터리가 실질적으로 수집한 파일 수(R)는 다른 리포지터리와 비교하여 낮은 편이다. 하지만 수집한 파일에 포함되어 있는 인용 및 참고문헌 정보와 링크를 함께 제공함으로써 리포지터리 크기(S)와 가시성(V)이 매우 높게 나타났음을 알 수 있다. 이들 리포지터리의 성공은 많은 자원을 수집하는 것이 어렵거나 부족한 경우, 다른 사이트와 연계체제를 구축함으로써 크기와 가시성을 향상시킨 좋은 사례이다.

또한 다양한 검색방식과 이용자 친화적인 인터페이스를 제공하는 것은 리포지터리 활성화에 매우 중요한 요소이다. TOP 10 리포지터리의

경우 모두 두 가지 이상의 검색 인터페이스를 제공한다. arXiv는 저장소와 논문번호를 아는 경우와 그렇지 않은 경우를 구분하여 두 가지 인터페이스를 제공하고 있으며, DTIC는 세 가지 검색 인터페이스인 DTIC Online, DTIC Online Access Controlled, DTIC Online Classified를 제공하고, ADS는 Basic Query Form 외 총 7가지 검색 인터페이스를 제공한다. DSpace@MIT는 회원에 따라 Core Service와 Premium Service로 구분하여 서비스를 이원화한다.

<표 3> Ranking Web of World Repositories 순위

종합 순위	이름	S	V	R	Sc
2	CiteSeerX	1	1	411	2
5	RePEc	2	6	140	5
6	ADS	3	4	539	1

4.5 보존정책

앞에서 밝힌바와 같이 리포지터리의 중요한 역할 중 하나는 디지털 콘텐츠의 장기보존이다. 리포지터리의 제출한 논문의 유지, 관리, 장기 보존, 그리고 지속적인 서비스에 대한 약속은 더 많은 이용자 참여를 유도할 수 있다.

대부분의 리포지터리에 일단 제출된 자료는 저자 요청 없이는 운영주체나 다른 연구자들이 삭제하거나 수정할 수 없다. 특히 arXiv나 CERN은 제출된 논문에 대한 장기적인 보존 정책을 다른 리포지터리보다 강력하게 명시하고 있다. arXiv는 저자의 요청이 있거나 초판, 수정판, 최종판이 제출되더라도 각각 별도의 레코드로 관리하고 이전의 것을 삭제하지 않는다. CERN의

경우에도 제출된 논문의 보존을 위해 저자가 요청하더라도 파일의 수정이나 삭제가 일부 제한된다.

이러한 리포지터리의 보존과 서비스 지속성에 대한 약속은 제출자에게 신뢰성을 주어 이용자의 활발한 참여를 유도하는데 도움이 된다. 하지만 국내의 리포지터리 사례들은 보존이나 서비스 지속에 대한 약속을 명시하고 있는 곳이 없다.

4.6 기술정책

리포지터리를 설계하거나 구축할 때 사용할 수 있는 다양한 소프트웨어들이 개발되어 있다. 자체 개발의 방식이 아닌 시중의 소프트웨어를 사용하면 소프트웨어 개발에 드는 비용과 노력을 줄일 수 있을 뿐만 아니라 많은 곳에서 사용하고 있는 기술을 채택함으로써 리포지터리간의 상호운용성도 확보할 수 있다는 장점이 있다.

Registry of Open Access Repositories에 등록되어 있는 1,853개의 리포지터리들이 사용하고 있는 개발 소프트웨어를 살펴보면, DSpace를 717개 기관에서 사용하고 있고 자체개발의 소프트웨어를 사용하는 곳이 442개, Eprints 350개, Bepress 101개의 순이다.

본 연구에서 분석한 성공적인 리포지터리의 경우 주제 리포지터리 5개와 기관 리포지터리 1개가 자체 개발 소프트웨어를 사용하는 것으로 나타났다. 기관 리포지터리 중 CERN과 MIT@DSpace은 자체 개발한 소프트웨어 각각을 유사 기관의 리포지터리들의 구축을 위해 배포하여 함께 사용하고 있는데, CERN은 15개 기관과 함께 CDS Invenio를 공유하고 있다. MIT가

개발한 DSpace는 가장 많은 수의 대학 및 연구소들이 기관 리포지터리 구축을 위해 사용하고 있고 우리나라의 S-Space와 KOASAS도 이에 해당된다.

4.7 저작권정책

리포지터리는 오픈액세스 기반의 학술정보 보존 및 확산을 목적으로 개발되었기 때문에 대부분의 리포지터리들이 제출된 연구 성과물들에 대해 이용자에게 제한없이 무료로 이용할 수 있게 한다. 하지만 현재 저작권 상황은 출판사나 학회를 통해 논문을 출판하면 저작권이 출판사로 이양되는 구조이기 때문에 이미 공식적인 출판을 거친 논문들은 저자가 직접 리포지터리에 제출하더라도 저작권 문제가 발생할 수 있다. 리포지터리에서 수집된 연구 성과물들에 대한 저작권을 이양받기 위해 저작권이 저자에게 있는 출판전 기사나 출판 중인 논문 위주로 저자 직접 제출을 통해 수집하고 있다. arXiv, SSRN, RePEc가 대표적인 예이다. 출판된 논문을 리포지터리에 아카이빙하는 경우에는 출판사의 저작권 위임서나 승인서를 함께 제출하도록 하고 있다. 우리나라의 S-Space와 KOASAS는 제출자가 자료의 제출전에 출판사나 저널의 저작권 정책을 확인 후에 제출하도록 하고 있다.

DSpace@MIT는 소속 교수들과 연구원들에게 출판사 출판일과 같은 날에 논문을 제출하도록 의무화하고 있다. 이는 연구 중인 논문의 경우, 지속적인 수정 과정이 남아 있는 상태이기 때문에 품질이 낮을 수 있으므로 최종 논문을 제출받아 보다 양질의 논문을 수집하고 동시에 저작권 문제를 해결하기 위한 전략이다.

또한 MIT에서는 출판사가 아카이빙을 허락하지 않는 경우 다른 출판사로 변경을 요구하는 보다 강력한 가이드라인을 제시하고 있다.

리포지터리에서 양질의 학술 논문을 저작권에 구애받지 않고 오픈 액세스 기반의 자유로운 유통을 위해서는 향후 출판사나 에이전시와 지속적인 논의와 협상이 필요할 것이다.

4.8 리포지터리간 운영정책 비교

리포지터리의 성공 사례들을 분석한 결과, 앞에서 기술한바와 같이 운영 전반에 걸쳐 공통적인 부분과 상이한 부분들이 분석되었고, 주제 리포지터리와 기관 리포지터리의 운영에서의 차이점 등이 도출되었다. 먼저 분석에 사용된 성공적인 리포지터리들이 포함하고 있는 운영정책의 주요 요소들을 정리해보면 <표 4>와

같다.

리포지터리간 운영정책의 강도나 상세 수준은 다양하지만 성공적인 리포지터리 운영을 위해 연구자들의 자발적인 참여와 이용자들에게 이용의 용이성을 제공하는 측면은 모두 유사하게 나타남을 알 수 있다.

성공 사례에 해당하는 리포지터리들은 모두 자국어와 병행하여 영어로 된 홈페이지 및 세부 섹션 페이지를 제공하고 있었고 제출물의 언어에는 특별한 제한이 없었다. 그리고 논문 제출자는 본인 등록 절차를 거쳐 제출이 가능하지만 이용의 용이성을 위해 기본적인 서비스 이용에는 어떤 이용 제한도 없었다. 또한 제출되는 논문들은 고유한 URL이나 식별자를 부여해서 접근의 용이성을 제공하고 있었다. 그 외에도 이용의 활성화를 위해 다양한 검색 서비스 및 알리미 서비스, 인용정보 서비스, 이용 통계 서비

<표 4> 리포지터리간 운영정책 비교

구 분		주제 리포지터리					기관 리포지터리				
		arXiv	Cite	SSRN	ADS	RePEc	DTIC	CERN	MIT	S-Space	KOASAS
운영 및 지원정책	운영권한 최소화	○	○	○	○	○					
	운영권한 강화						○	○	○		
수집정책	직접제출	○					○	○	○	○	○
	직접제출 + 로봇		○	○	○	○	○	○			
관리정책	최소의 품질관리	○	○	○	○						
	자료포맷 제한	○	○	○	○	○			○		
	고유식별자 부여	○	○	○	○	○	○	○	○	○	○
	영어 + 자국어	○	○	○	○	○	○	○	○	○	○
배포 및 이용활성화 정책	이용제한 없음	○	○	○	○	○	○	○	○		○
	메일링 서비스	알리미	○	○	○	○	○	○	○		
		인용정보		○		○	○		○		
		이용통계					○		○		
보존정책	명문화	○		○				○			
기술정책	상호운용성 지원							○	○	○	○
저작권정책	배포전 논문 수집	○		○		○			○		

스 등을 제공하는 것으로 나타났다.

4.8.1 주제 리포지터리와 기관 리포지터리
리포지터리의 성공 사례에서 나타나는 공통적인 운영 정책 외에 리포지터리의 큰 범주인 주제 리포지터리와 기관 리포지터리간의 차이가 있음을 발견할 수 있었다. 주제 리포지터리들의 경우에는 운영 주체의 권한을 최소화하는 반면 기관 리포지터리들은 제출에 대한 의무조항이나 기관내의 위원회를 통해 지원을 받는 형식으로 보다 운영주체의 권한이 강화된 방식이었다. 다량의 콘텐츠 축적을 위해 주제 리포지터리들은 직접 제출뿐만 아니라 로봇, 기관 간 협력, 기존 데이터베이스 활용 등의 복수적인 방식을 통해 콘텐츠를 확보하는 반면 기관 리포지터리들은 저자의 직접 제출 방식을 지향하고 있으며 그렇지 못한 경우, 특히 CERN은 다른 리포지터리를 통해 콘텐츠를 확보하는 것으로 나타났다. 리포지터리를 구축하기 위한 소프트웨어는 비교적 구축 범위가 제한되어 있는 기관 리포지터리의 경우에는 상호운용성을 위해 공동의 소프트웨어를 사용하지만 주제 리포지터리의 경우에는 모두 자체개발 소프트웨어를 사용하는 것으로 나타났다.

즉, 범 기관적인 주제 리포지터리는 자발적인 정보 개방과 공유의 도구라는 역할에 초점이 맞추어져 있는 반면 기관 리포지터리는 소속 연구자들의 성과물 관리 및 보존, 기관의 위상 정립, 상위 기관의 사명과 목표에 부합 등의 좀 더 다양한 역할이 중요시 되고 있는 것으로 나타났다. 이러한 이유에서 리포지터리에 제출하는 자료의 품질 관리에 대해서도 주제 리포지터리는 최소의 품질 관리를 지향하여 되도록

많은 연구자들의 참여를 유도하는 반면 기관 리포지터리의 경우에는 기관내 제출자의 신분 에 따라 제출의 가능 여부를 결정하는 등의 품질 관리 규정들이 강화되어 있다.

4.8.2 국내와 국외의 리포지터리 성공 사례
S-Space와 KOASAS를 제외하고는 Ranking Web of World Repositories의 800위에 포함된 국내 리포지터리가 없다는 것과 그나마 포함되어 있는 S-Space와 KOASAS의 경우에는 모두 500위 밖에 밀려 있다는 것을 알 수 있었다. 국내에 기반을 두고 있는 주제 리포지터리는 거의 없고 기관 리포지터리의 활성화도 미진하여 국내 리포지터리의 성공적인 구축과 운영을 위한 전략을 모색하고자 Top 10에 포함되어 있는 국외의 성공 사례와 500위권의 S-Space와 KOASAS의 운영 정책을 추가로 비교해 보았다.

전반적인 운영정책에서 국외의 성공 사례와 달리 국내의 S-Space와 KOASAS는 보다 구체적이고 세분화된 리포지터리 운영 정책이나 가이드라인을 포함하고 있지 않다. 국외의 기관 리포지터리들이 포함하고 있는 운영주체의 강력한 권한에 대한 문항, 제출에 대한 의무조항, 자료 포맷에 대한 지원, 보존의 지속성에 대한 문항, 저작권 정책에 대한 보다 강력한 대응을 위한 문항 등이 S-Space와 KOASAS에는 없다.

특히 저자의 자율적인 직접 제출에 전적으로 의존하고 있음에도 불구하고 기관내 소속 연구자들의 연구 성과물 제출에 대한 적극적이고 의무적인 문항들이 없으며 제출되지 않은 연구 성과물의 수집 방법에 대한 대안도 제시되어

있지 않다.

또한 리포지터리의 지속성 및 리포지터리에 제출된 자료들의 보존에 대한 신뢰성이 리포지터리 활성화를 이끄는 중요한 요인임에도 불구하고 이러한 보존 정책에 대한 명문화된 정책을 포함하고 있지 않다. 마찬가지로 국내의 리포지터리들은 자료 제출시 저작권에 대한 확인 절차에 대한 설명만을 포함하고 있고 저작권 정책에 구속받지 않는 출판전 논문 수집에 대한 적극성과 저작권 정책에 대응할 수 있는 다양한 방법들에 대한 적극적인 방안들을 포함하고 있지 않다.

무엇보다 국내 리포지터리는 배포 및 이용 활성화 정책이 미비하다. 제출된 논문의 상태만 확인할 수 있지 이용자들의 참여를 유도할 수 있는 인용정보에 대한 실시간 알리미 서비스, 이용통계서비스 등을 제공하지 않고 국외의 성공적인 리포지터리와 비교하여 다양한 검색서비스와 경로를 제공하고 있지 않은 것으로 나타났다. S-Space와 KOASAS 모두 Ranking Web of World Repositories의 평가 기준 네 가지에서 특히 가시성(각각 781위와 720위)이 상대적으로 매우 낮게 나타났다. 그것은 기관의 리포지터리 사이트 외에 검색엔진 Google에서만 검색을 지원하고 있기 때문이다. 국외의 성공적인 리포지터리들은 세계의 4대 검색엔진인 Google, Yahoo, Live Search, 그리고 Exalead에 모두 노출되어 있다. 리포지터리의 가시성을 높이는 것이 리포지터리 활성화를 위한 최대의 관건인 만큼 대규모의 검색엔진들에게 노출될 수 있도록 노력하는 것이 필요하다.

5. 결론

학술정보 유통의 새로운 패러다임인 오픈액세스 기반 지식정보 유통을 실현하기 위해서는 리포지터리의 역할이 무엇보다 중요하다. 해외 많은 대학이나 연구소에서는 이미 리포지터를 활발하게 구축·운영하고 있으며 향후 급속하게 증가할 전망이다. 하지만 국내의 경우에는 여전히 리포지터리의 중요성에 대한 인식이 부족하고 미비하다.

이에 본 연구에서는 성공적인 국내 리포지터리 구축·운영 및 활성화 전략을 모색하기 위해 성공적인 것으로 평가되는 해외 리포지터리들의 사례를 운영 및 지원, 수집, 관리, 배포 및 이용활성화, 장기보존, 기술, 저작권 정책의 측면에서 조사·분석하였다. 주제 리포지터리와 기관 리포지터리 운영에 있어서의 차이점과 국내와 국외 성공 사례의 차이점도 분석해보았다. 사례분석에 사용한 성공적인 리포지터리는 스페인의 CSIC에서 TOP 10으로 선정된 것이며 이들 중 사이트를 통해 정책 전반을 밝히고 있는 8개로 크게 주제와 기관 리포지터리로 구분하였다. 추가로 이 사이트에 랭크되어 있는 국내 리포지터리 2개도 포함하였다.

본 연구에서 도출된 성공적인 리포지터리를 구축·운영하기 위한 전략과 시사점들은 모든 리포지터리에 일괄적으로 적용하기에는 무리가 있을 것이다. 크게는 주제와 기관 리포지터리 구축의 목적과 역할이 달라서 운영에 있어 차이점이 있는 것과 마찬가지로 기관에서 정의한 리포지터리의 역할과 특성이 매우 다양함으로 이에 따른 성공적인 운영정책이나 전략은 세분화될수록 더 다양해질 것이다. 하지만 본

명한 것은 리포지터리의 근본적인 목적이 연구자의 자발적인 참여와 이를 통한 개방적인 정보공유 환경을 조성하는 것이다. 그래서 리포

지터리의 주요 연구자나 이용자에게 너무 많은 규제나 제한은 성공적인 운영의 큰 장애 요인이 될 수 있다.

참 고 문 헌

- 황혜경, 김혜선, 최선희. 2004. 오픈액세스기반 지식정보저장소 구축에 관한 연구. 『한국비블리아학회지』, 15(1): 91-116.
- 황혜경, 이지연. 2009. 오픈액세스기반 리포지터리 성공에 미치는 요인 분석. 『정보관리학회지』, 26(4): 35-56.
- Crow, Raym. 2002. "The Case for Institutional Repositories: A SPARC Position Paper." [online]. [cited 2010.4.23]. <http://scholarship.utm.edu/20/1/SPARC_102.pdf>.
- Johnson, Richard K. 2002. "Institutional Repositories: Partnering with Faculty to Enhance Scholarly Communication." *D-Lib Magazine*, 8(11). [online]. [cited 2010.4.20]. <<http://www.dlib.org/dlib/november02/johnson/11johnson.html>>.
- Lercher, Aaron. 2008. "A Survey of Attitudes About Digital Repositories Among Faculty at Louisiana State University at Baton Rouge." *The Journal of Academic Librarianship*, 34(5): 408-415.
- Lynch, C. A. 2003. "Institutional Repositories: Essential Infrastructure for Scholarship in the Digital Age." [online]. [cited 2010.6.9]. <<http://www.arl.org/newsltr//226/ir.html>>.
- Proudman, Vanessa. 2007. "CERN Document Server: Institutional Repository and Service. CERN, Geneva, Switzerland." [online]. [cited 2010.11.4]. <<http://arno.uvt.nl/show.cgi?fid=68180>>.
- Robinson, Michael. 2009. "Promoting the Visibility of Educational Research through an Institutional Repository." *Serials Review*, 35(3): 133-137.
- Rumsey, Sally. 2006. "The Purpose of Institutional Repositories in UK Higher Education: A Repository Manager's View." *International Journal of Information Management*, 26: 181-186.
- Shearer, Kathleen. 2002. "Institutional Repositories: Towards the Identification of Critical Success Factors." *The Canadian Journal of Information and Library Science*, 27(3): 250-263.
- Westell, Mary. 2006. "Institutional Repository

- ries: Proposed Indicators of Success.” *Library Hi Tech*, 24(2): 211-226.
- Xia, Jingfeng. 2007. “Assessment of Self-Archiving in Institutional Repositories: Across Disciplines.” *The Journal of Academic Librarianship*, 33(6): 647-654.
- Xia, Jingfeng and Sun Li. 2007a. “Assessment of Self-Archiving in Institutional Repositories: Depositorship and Full-Text Availability.” *Serials Review*, 33: 14-21.
- Xia, Jingfeng and Sun Li. 2007b. “Factors to Assess Self-Archiving in Institutional Repositories.” *Serials Review*, 33: 73-80.
- 〈참고 사이트〉
- arXiv.org e-Print Archive. <<http://arxiv.org>>.
- CERN Document Server.
<<http://cdsweb.cern.ch>>.
- CiteSeerX. <<http://citesserx.ist.psu.edu/>>.
- Defense Technical Information Center.
<<http://www.dtic.mil/dtic/>>.
- DSpace@MIT. <<http://dspace.mit.edu/>>.
- KOASAS. <<http://koasas.kaist.ac.kr>>.
- Ranking Web of World Repositories.
<<http://repositories.webometrics.info/>>.
- Registry of Open Access Repositories.
<<http://roar.eprints.org/>>.
- Research Papers in Economics.
<<http://repec.org>>.
- Smithsonian/NASA Astrophysics Data System. <<http://adsabs.harvard.edu/>>.
- S-Space. <<http://s-space.snu.ac.kr>>.
- SSRN. <<http://ssrn.com>>.