

디지털 자원의 아카이빙 우선순위 결정 방법에 관한 연구

A Study on Prioritizing Digital Archiving Resources

김광영(Kwang-Young Kim)*, 곽승진(Seung-Jin Kwak)**

【초 록】

도서관이나 학술정보센터에서 디지털 아카이빙의 필요성은 증대되고 있지만 모든 디지털 자원들을 아카이빙하기에는 전문 인력과 예산이 많이 부족한 상황이다. 따라서 자원의 보조가치와 이용가치 등의 요소를 고려한 우선순위에 따라 순차적으로 아카이빙 처리를 할 필요가 있다. 본 연구에서는 한국과학기술정보연구원(KISTI)이 보유하고 있는 디지털 자원을 중심으로 어떤 콘텐츠를 우선적으로 아카이빙을 해야 하는지 이용가치를 중심으로 조사하였다. 또한 이것을 기반으로 자동으로 시스템 분석을 수행하고 적용할 수 있는 우선순위 평가 방안을 제안하였다.

【키워드】

디지털 자원, 디지털 콘텐츠, 디지털 아카이빙, 우선순위 결정, 우선순위 평가, 한국과학기술정보연구원

【Abstract】

A necessity increases but since a many manpower and the budgets are insufficient for archiving all digital resources the digital archiving at the science information center or the digital library has to be successively processed according to the priority. Therefore, in this research, whether it has to archive among digital resources which Korea Institute Science and Technology Information(KISTI) hold around any kind of contents or not makes analysis and valuation around the value for use. The priority evaluation plan that can become the system analysis and reflection based on this automatically tries to be proposed.

【Keywords】

Digital Resources, Digital Content, Digital Archive, Priority Decision, Priority Evaluation, KISTI

1. 서론

1.1 연구 필요성 및 목적

오늘날 우리는 디지털 정보의 홍수 속에서 살고 있다. 디지털 정보가 기하급수적으로 늘어나고 있지만 도서관이나 보존 전문기관의 관리가 이루어지기 전에 급속한 폐기와 망실 또한 일어난다. 최근 각국에서는 정부와 민간 차원에서 디지털 자원의 생산뿐만 아니라 항구적 보존을 위한 시범시스템 구축 등 다양한 노력들을 기울이고 있다.

대부분 디지털 자원들은 매체의 특성상 변조 가능성이 높고 출판사 등 생산자의 사정에 따라서 폐기되는 등 인쇄자료에 비하여 휘발성이 매우 높다. 디지털 자원을 아카이빙하기 위해서는 확고한 아카이빙 정책, 전문 인력과 예산, 관련 기술개발 등의 지속적인 연구가 필요하다. 각 기관이 가지고 있는 모든 디지털 자원들을 아카이빙하기에는 인력과 예산이 부족하기 때문에 우선적으로 자원을 선별하여 순차적으로 아카이빙을 수행하고 국가의 대표적인 기관간의 협력을 통한 아카이빙이 요구된다.

우리나라의 경우 디지털 자료에 대한 아카이빙의 중요성은 점차적으로 인식하고 있으나 본격적으로 추진되고 있다고 볼 수는 없다. 유네스코 한국위원회와 정보트러스트센터가 주최하고 국립중앙도서관이 후원한 디지털 유산 보존을 위한 포럼에서 보존에 대한 인식 제고를

* 한국과학기술정보연구원 정보기술연구실 선임연구원(kykim@kisti.re.kr) (제1저자)

** 충남대학교 문헌정보학과 부교수(sjkwak@cnu.ac.kr) (교신저자)

논문접수일자 : 2010년 11월 5일 논문심사일자 : 2010년 11월 30일 게재확정일자 : 2010년 12월 10일

촉구하는 논의가 이루어진 바 있고, 국립중앙도서관에서 온라인 디지털자원 보존사업을 추진하고 있으나 아직 국가지식자원 관리 차원에서 보존의 문제를 인식하고 있다는 흔적은 엿보이지 않는다(한국과학기술정보연구원 2005).

정부는 국가의 재정적 지원으로 생산된 공공지식정보자원을 체계적으로 수집, 유통하기 위해 책임 기관을 지정하고 있다. 그 기관으로는 한국과학기술정보연구원(KISTI), 한국교육학술정보연구원(KERIS), 국립중앙도서관 등이 있다. 특히 대표적인 과학기술분야의 종합정보센터인 KISTI는 교육과학기술부 산하 기관들이 생산한 자원을 대상으로 지식정보를 구축하도록 지정되었다.

KISTI는 학술 논문, 특허, 연구보고서, 동향분석, 산업표준, 사실정보 등의 다양한 디지털 자원을 생산에서부터 수집과 서비스, 유통하는 대표적인 기관이다. 2010년 1월 기준으로 디지털 자원의 구축현황을 보면 약 9천만 건 이상이다. KISTI의 디지털 자원의 항구적인 보존과 미래 이용자를 위한 접근을 보장하기 위해서 모든 디지털 자원을 아카이빙하기 위해서는 많은 전문적인 인력과 지속적인 예산의 뒷받침이 필요하다. 연구자에 따라서 어떤 디지털 자원을 우선적으로 아카이빙해야 하

는지에 대한 주장은 다양하다. 즉 수량이 가장 많은 학술자료에 대해서 우선순위를 두고 아카이빙을 요구하는 사람들도 있고 해외 자료보다는 국내자료를 우선적으로 해야 한다고 주장하는 사람들도 있다. KISTI의 디지털 자원들에 대한 아카이빙 우선순위에 대한 연구 결과에서는 보존가치와 이용가치가 가장 높은 순위로 중요도가 나타났다(KISTI 2010).

본 연구의 목적은 현재 KISTI에서 서비스하고 있는 많은 디지털 자원들의 보존가치 보다는 이용가치 측면에서 우선적으로 선별하여 아카이빙해야 할 디지털 자원이 무엇인지에 대하여 분석한다. 또한 이런 요소들이 우선순위 평가 시스템에서 자동으로 반영될 수 있는 우선순위 평가 방안을 제안하는 것이다.

1.2 연구 범위 및 방법

본 연구는 현재 KISTI가 운영하는 NDSL의 콘텐츠들 중에서 국내학술지, 해외학술지, 연구보고서, 동향분석 정보 및 특허를 대상으로 아카이빙 우선순위를 조사한다. 2010년 1월 기준 KISTI의 디지털 콘텐츠 구축 현황은 <표 1>과 같다.

<표 1> KISTI의 디지털 콘텐츠 구축 현황

유형	구분	세부구분	수록내용	수록기간	구축건수
논문	학술지	국내학술지	국내논문지, 협회지, 기관지, 동향지 수록논문	1948-현재	681,211
		해외학술지(영미권)	해외 과학기술분야 핵심 학술지 수록논문	1666-현재	42,422,251
		해외학술지(중·일)	KISTI 보유 중국, 일본 학술지 수록논문	2003-현재	3,181,435
	학술회의	국내학술회의	국내 학술대회 발표논문	1972-현재	278,214
		해외학술회의	해외 학술대회 발표논문	1993-현재	7,397,735
		학위논문	국내 석박사 학위논문	1945-현재	1,280,680
소 계					55,241,526
연구보고서		국가연구개발보고서	정부산하 기관의 재정지원을 받아 수행된 연구개발보고서	1983-현재	115,994
		미국연구개발보고서	미국정부의 재정지원을 받아 수행된 연구개발보고서	1995-현재	156,918
	소 계				272,912
동향분석		분석리포트	KISTI 산업, 기술정보의 고급 분석보고서	최신	23,321
		글로벌 동향브리핑	해외 과학기술동향 정보서비스	최신	133,873
		과학기술 정책동향	국내외 과학기술 정책동향(S&T GPS)	최신	16,233
	소 계				173,427
특허		한국특허	한국 출원, 공개/등록, 실용신안, 의장특허	1983-현재	3,245,340
		미국특허	미국 공개/등록 특허	1976-현재	6,351,039
		일본특허	일본 공개특허	1976-현재	8,550,932
		유럽특허	유럽 공개특허	1976-현재	2,222,765
		국제특허	WIPO PCT 특허	1976-현재	1,887,105
	소 계				22,257,181

※ 출처 : KISTI 내부 자료.

『지식정보자원관리법』의 규정에 의하면 국가적으로 관리해야할 지식정보의 대상 분야를 보존가치, 이용가치, 내재적 가치로 규정한다. 보존가치는 역사성, 문화적 정체성, 민족성 등 한 국가 또는 민족의 고유성을 의미한다. 보존가치가 있는 지식정보자원은 모든 국가에서 일차적으로 우선 관리대상으로 삼는 지식정보자원이다. 이용가치는 이용에 대한 관심, 이용가능성, 이용을 통한 부가가치성, 이용자 범위를 의미한다. 이용자의 관심 또는 인식에 따라 변이의 폭이 상대적으로 클 수 있다. 내재적 가치는 대상지식정보자원이 본질적으로 함유하고 있는 절대적 의미의 질적 특성을 의미한다. 보존가치, 이용가치, 내재적 가치 등은 상황에 따라서 상호 밀접하게 관련을 맺을 가능성도 높다. 『지식정보자원관리법』의 이러한 기준은 아카이빙 대상의 선정에도 상당한 의미를 갖는다.

이와 같이 아카이빙 대상 콘텐츠의 우선순위 결정을 위한 평가기준의 고려사항들은 상당히 많음을 알 수 있다. 2005년 KISTI의 연구에서 디지털 아카이빙 대상의 선정 기준으로 ① 생산조직 및 기관의 중요도, ② 이용가치, ③ 정보의 고유성, ④ 국가자원으로서의 보존가치 등으로 제시되었다(KISTI 2005).

아카이빙 대상 콘텐츠 선정시 평가기준과 고려사항을 비교한 연구를 정리하면 보존가치, 이용가치, 중요성, 내재적 가치, 보존타당성과 보존비용을 고려할 것으로 제시하였다. 콘텐츠 유형별 중요도 측정 및 참여 연구자들의 중요도 인지과정을 심층적으로 분석하기 위한 설문조사 및 포커스그룹 토의를 통한 결과를 보면 보존가치와 이용가치가 가장 높은 순위로 중요도가 나타났다(KISTI 2010).

아카이빙 우선순위 설정을 위해서는 여러 연구에서 제

시한 다양한 요소들을 고려할 필요가 있다. 하지만 기관의 특성상 이런 다양한 요소들을 모두 적용한 모델을 정립하기에 경제적, 기술적, 법적측면의 어려운 점들이 많다.

본 연구에서는 현재 이용가치가 높은 자원은 향후에도 높은 이용 잠재성을 가질 것이라고 가정하고, KISTI의 아카이빙 대상 자원 선정을 위한 우선순위 결정에서 보존가치보다는 이용가치를 중심으로 분석한다.

연구 방법으로는 먼저 KISTI에서 보유하고 있는 디지털 자원과 실제 사용자들이 얼마나 이용하고 있는지 현황을 조사하고, 보유량과 이용률을 산출하여 대상 콘텐츠들에 대해서 평가를 실시한다. 이것을 기반으로 디지털 자원의 아카이빙을 위한 우선순위 결정을 위한 기초 자료로 활용하여 우선순위 평가 방안을 제시한다.

2. 선행 연구

2.1 국내 아카이빙 현황

국내의 경우, 디지털 콘텐츠에 대한 아카이빙은 시작단계이다. 현재 국회도서관, KERIS, KISTI 등의 국가기관과 북도피아, DBPIA 등의 민간기업체에서 디지털 콘텐츠를 저장하고 있고 시범적으로 디지털 아카이빙 정책을 구상하는 수준이다. 최호남(2005)은 우리나라 디지털 아카이브는 장기 보존 및 접근에 대한 체계적 접근이 없는 상태로 진정한 의미의 아카이브로 보기는 어려운 실정이라고 하였다. 국내 각 기관별 아카이빙 정책 비교는 <표 2>와 같다. 국립중앙도서관, 국가기록원 등은 장기보존 전략을 가지고 있으며 최근 KISTI에서도 디지털 자원의 항구적인 보존과 접근을 위한 정책을 수립하고 있다.

<표 2> 각 기관별 아카이빙 정책 비교

구 분	국립중앙도서관	국가기록원	국회도서관	KERIS	한국학술정보	DBPIA	KISTI
							학회마을
주제 분야	전 주제	전 주제	전 주제	전 주제	전 주제	전 주제	과학기술
목적	공공이익	공공이익	공공이익	공공이익	상업적	상업적	공공이익
대상 콘텐츠	국내의 디지털 장서, 국내의 학술문화자원, 디지털 문화유산 등	정부 간행물	석박사논문, 국내학술지, 정부간행물	국내 석박사 학위논문	국내 학/협회지	국내 학/협회지	국내 학/협회지
장기보존 전략 유무	○	○	X	X	X	X	○
인터넷 서비스	○	○	○	○	○	○	○
기관 성격	국가기관	국가기관	국가기관	출연기관	민간기업	민간기업	출연기관

2.2 국외 아카이빙 현황

2.2.1 학술지 아카이빙 사례

학술지는 심사제(Peer-review)를 통해 검증된 학술 연구정보를 관련 분야 연구자들에게 공식적으로 제공하는 가장 신뢰할 수 있는 핵심정보자원으로 평가되고 있는 중요한 디지털 정보 자원이다.

네덜란드 국가도서관(Koninklijke Bibliotheek/National Library of the Netherlands, KB)에서는 e-Depot을 중심으로 매체노후화 및 소프트웨어와 하드웨어 플랫폼의 빠른 진보에 대비하여 디지털 객체에 대한 장기적인 접근을 보장하는 것을 목적으로 자국의 디지털 납본자료, 웹 아카이브, 디지털화 자료는 물론 전 세계 디지털발간물을 아카이빙하고 있다. 2003년부터 2008년 12월 현재까지 e-Depot이 수집한 1,200만 건 이상의 학술지 대부분이 Elsevier, Springer 등의 국제적 출판사의 발간물들이다.

Portico는 2002년 앤드류멜론재단(Andrew W. Mellon Foundation)의 재정지원으로 JSTOR가 시작한 Electronic-Archive Initiative에서 출발하였다. 전자학술지의 장기적 아카이브를 제공하는 비영리서비스로 2005년부터 새로운 전자 아카이빙 서비스가 시작되었고 앤드류멜론재단은 물론이고, 최근에는 미국 의회도서관(Library of Congress, LC)의 재정지원을 받고 있고 있다. 참여 출판사가 제공하는 전자형태로만 발간되는(born digital) 전자학술지, 인쇄 및 전자형태가 병행 발간되는 학술지, 원래 인쇄본을 디지털화한 학술지 모두를 아카이빙 대상에 포함시키고 있다.

LOCKSS(Lots of Copies Keep Stuff Safe)는 스탠포드대학 도서관(Stanford University Libraries)에 본부를 두고 있는 국제적 프로그램으로 도서관이 자관 소유의 디지털콘텐츠를 저렴한 비용으로 쉽게 수집·보존할 수 있도록 디지털보존 툴(공개 소프트웨어)과 각종 지원 서비스를 제공하고 있다. LOCKSS 회원도서관은 현재 약 400개 학술출판사가 발간하는 전자학술지 2,700종의 보존권한을 획득하였다. 아시아/태평양지역에서는 중국, 뉴질랜드, 싱가포르 도서관이 회원으로 참여하고 있고 있다.

Trusted Digital Repository(TDR, NRC-CISTI)는 캐나다 국립과학기술정보연구원(National Research Council-Canada Institute for Scientific and Technical Information, NRC-CISTI)에서 구축하는 TDR의 사명은 현

재와 미래에 신뢰할 수 있고 장기적으로 접근할 수 있도록 관리된 디지털 정보자원을 제공하는 것으로 TDR 콘텐츠는 향후 검색창을 통하여 제공할 예정이다. TDR 연계에 대한 네트워크 수행능력과 전자학술지의 접근성을 평가하고자 Federal Science eLibrary 파일럿 프로젝트를 수행했다. 이 프로젝트를 통하여 NRC-CISTI의 infostructure가 신뢰성 있는 플랫폼이라는 것과 캐나다 정부 연구원, 정책분석가 및 정책결정자가 신속하게 STM 분야의 핵심 전자학술지에 접근할 수 있어서 연구에 도움이 되었다는 긍정적 평가를 얻었다.

Journal@rchive는 일본 과학기술진흥기구(科學技術振興機構, Japan Science and Technology Agency, JST)에서 구축한 Journal@rchive는 과학기술정보 발신·유통 시스템(J-STAGE)의 아카이브 사이트로 2005년부터 실시하고 있는 전자 아카이브 사업이다. 일본의 연구 성과에 대한 국제 발신 능력 강화와 중요한 지적 재산의 보존을 목적으로 한다. 일본 자국내 학회회가 발간하는 학술지를 창간호부터 전자화하여 공개함으로써 메이지 시대부터 일본의 연구 성과를 열람할 수 있고, 세계에 영향을 준 저명한 논문도 많이 수록하고 있다. 전자 아카이브 대상지는 일본학술회의 등 관련 기관의 협력으로 선정하고 있으며, 자연과학분야뿐만 아니라 인문, 사회과학분야 학술지도 대상으로 함으로써 2005년 74개, 2006년 65개, 2007년 58개, 2008년 181개, 2009년 266개가 선정되었다. 2010년도 이후에도 추가로 선정하여 범위를 확대해 나갈 예정이다.

2.2.2 연구보고서 아카이빙 사례

연구보고서 또는 기술보고서는 산업계, 학계, 연구계에서 생산되는 과학기술분야의 최신 연구경향과 성과를 알 수 있는 유용한 수단이고 특히 관련분야의 전문가에게 가치 있는 정보를 제공하는 중요한 정보자원이다.

Technical Reports Archives and Image Library(TRAIL)는 미국의 애리조나대학(University of Arizona)이 주도하는 Greater Western Library Alliance(GWLA)는 Center for Research Libraries(CRL) 등과 협력하여 1975년 이전에 발행된 미국 연방정부의 기술보고서를 수집하고 디지털화하여 아카이브를 구축한 후 자유롭게 활용시키는 것을 목적으로 하고 있다.

Information Bridge는 미국 에너지부(Department of Energy, DOE)가 자체적으로 또는 연구비를 지원하여 생산한 연구보고서 전문(full-text) 및 서지정보 약 21만

건을 일반에게 무료로 제공한다. 주제 분야는 물리학, 화학, 재료학, 생물학, 환경학, 에너지기술, 컴퓨터 및 정보학, 재생에너지 등이다. 주로 1991년 이후에 생산된 것이지만 그 이전에 생산된 연구보고서의 디지털화 작업이 지속적으로 진행되고 있다.

NASA Technical Reports Server(NTRS)는 NASA의 기술문헌을 학생, 교사, 연구자, 일반인이 이용할 수 있도록 서비스를 제공한다. NASA가 자체 생산하거나 재정지원으로 생산되는 과학기술정보가 늘어남에 따라 계속 증가하고 있고 단일 인터페이스를 통해 서로 다른 3가지 컬렉션(NACA Collection, NASA Collection, NIX Collection)에 포함된 정보를 검색하는 것이 가능하다. NTRS에서 접근할 수 있는 정보 유형은 학술회의 논문, 이미지, 학술지 논문, 사진, 동영상, 특히, 연구보고서, 기술비디오 등 다양하고 현재 NTRS에서 약 50만 건의 항공우주 관련 자료의 서지, 온라인 원문 9만 건, 그리고 11만 1,000건의 이미지와 비디오 등을 이용할 수 있다. 연구보고서, 논문, 연구노트 등은 대부분 PDF 파일이며, 이미지, 사진, 동영상과 비디오의 포맷은 JPEG, GIF 및 QuickTime으로 저장하여 온라인으로 이용이 가능하다.

3. 디지털 아카이빙 대상 선정을 위한 평가 기준

디지털 아카이빙 대상 자원 선정을 위한 평가 기준과 고려사항은 다양한 관점에서 연구되어야 한다. 설문원(2005)은 생산조직 및 기관의 중요도, 이용가치, 정보의 고유성, 국가자원으로서의 보존가치 등으로 구체적으로 설정할 필요가 있다고 제시하였다. 아카이빙 대상 콘텐츠 선정시 평가기준과 고려사항을 비교, 정리하면 <표 3>과 같다(KISTI 2010).

보존가치는 아카이빙 대상 콘텐츠의 국가자원으로서의 보존가치를 의미하며 역사성, 문화적 정체성, 정보의 고유성 등을 고려하여 결정할 것을 제안하였다. 이용가치는 콘텐츠의 현재 이용가치는 물론 미래의 이용 잠재성을 함께 고려하여야 하며 접근성, 가용성, 이용을 통한 부가가치성, 그리고 이용범위 등을 고려할 것을 제시하였다. 중요성은 해당 콘텐츠의 생산조직 및 기관의 중요

도(즉, 정보출처의 중요성)와 함께 진본성 여부, 핵심적 증거 여부 등을 고려하는 평가기준으로 선정되었다. 내재적 가치의 경우 콘텐츠 자체가 지닌 절대적 질적 특성을 의미하는데 또한 자료의 무결성, 완결성을 내재적 가치를 결정하는 고려사항으로 포함하였다. 보존타당성은 보존가치와 구별되는 평가기준으로 해당 콘텐츠에 대한 저작권 소유여부, 유지상의 위험정도를 고려하여야 하겠다. 보존가치는 낮지만 영구손실의 위험이 큰 경우 장기적인 아카이빙의 필요성이 높다고 할 수 있겠다. 마지막으로, 보존비용의 경우 아카이빙에 필요한 인적·물적 비용 및 중장기 유지, 관리, 서비스 비용을 고려하여야 할 것을 제시하였다.

<표 3> 아카이빙 대상 선정시 평가기준 및 고려사항

평가기준	고려사항
보존가치	- 국가자원으로서의 보존가치 - 역사성, 문화적 정체성 - 정보의 고유성
이용가치	- 접근성/가용성 - 연구 활용도 - 이용에 대한 관심, 이용가능성, 이용 잠재성 - 이용을 통한 부가가치성 - 이용범위(이용자, 주제 분야, 정보수명 등) - 효율성, 보완성(기록의 관계와 종속성)
중요성	- 생산조직 및 기관의 중요도: 정보출처의 중요성 - 진본성 - 핵심적 증거
내재적 가치	- 절대적 질적 특성 - 무결성, 완결성
보존타당성	- 저작권 소유여부 - 유지상의 위험정도
보존비용	- 아카이빙에 필요한 인적, 물적 비용 - 중장기 유지, 관리, 서비스 비용

※ 출처: 지식정보자원관리법, 한국전산원, NDIIPP¹⁾, JISC²⁾, NARA, 캐나다 국립도서관 및 기록관, Ayris, 설문원의 등

<표 3>은 디지털 아카이빙 대상 선정시 적용할 수 있는 여섯 가지의 평가 기준(보존가치, 이용가치, 중요성, 내재적 가치, 보존타당성, 보존비용)을 일차적으로 선정하였다. 과학기술정보 DB 구축현황에서 제시된 바와 같이 국내 학술지, 해외 학술지, 국내 학술회의 자료, 해외 학술회의 자료, 학위논문, 국내 연구개발 보고서, 해외 연구개발 보고서, 국내 특허자료, 국외 특허자료, 동향분

1) <<http://www.digitalpreservation.gov>>.

2) <<http://www.jisc.ac.uk>>.

석, 산업표준, 사실정보, 웹 콘텐츠로 구분하여 연구를 수행하였다.

본 연구에서의 평가 방법은 보존가치보다는 이용가치를 중심으로 평가를 하며, 이용가치 중에서도 보유량과 이용률을 중심으로 대상 콘텐츠들에 대한 평가 요소로 활용 한다.

4. 평가 방법

디지털 아카이빙 대상의 선정을 위한 평가 기준에서 가장 높은 순위인 보존가치는 아카이빙 대상 콘텐츠의 국가자원으로서의 보존가치가 있다. 이용가치는 콘텐츠의 현재 이용가치는 물론 미래의 이용 잠재성을 함께 고려하여야 하며 접근성, 가용성, 이용을 통한 부가가치성, 그리고 이용범위 등을 고려해야 한다.

본 연구에서는 현재 이용가치가 높은 자원은 향후에도 높은 이용 잠재성을 가질 것이라는 가정에서 출발한다. 따라서 단순히 사용자의 이용률만을 분석하는 방법보다는 보유하고 있는 자원의 보유량에 대해서 얼마나 많은 사용자들이 이용하고 있는지를 분석한다. 이 결과는 디지털 아카이빙 대상 자원의 결정을 위한 기초 자료로 활용될 수 있을 것이다.

〈표 4〉는 2009년도 KISTI 정보 자원에 대한 이용 현황을 나타낸 것이다.

〈표 4〉 2009년도 KISTI 정보 자원 이용 현황

DB별	상세보기(횟수)	비율(%)
해외학술논문	1,214,137	28.81
국내학술논문	703,898	16.70
국내학위논문	47,662	1.13
동향정보	1,432,860	34.00
연구보고서	345,723	8.20
특허	284,824	6.76
표준	9,626	0.23
기타	175,495	4.16
합계	4,214,225	100.00

〈표 4〉에서와 같이 단순히 사용자의 이용 현황으로 분석할 경우에는 동향 정보, 해외 학술저널, 국내 학술저널, 연구 보고서, 특허 등의 순서로 아카이빙 우선순위를 결정할 수가 있다.

전체 데이터 보유량에 비해서 얼마나 많은 사용자들이 이용하는가를 중심으로 분석하기 위해서 전체 데이터 보유량에 대한 각 자원들의 보유비율을 조사하였다. 그 결과는 〈표 5〉와 같다. 〈표 5〉에서와 같이 국내 학술저널은 〈표 1〉에서 국내학술지, 국내학술회의 및 학위논문으로 구성을 한 것이다.

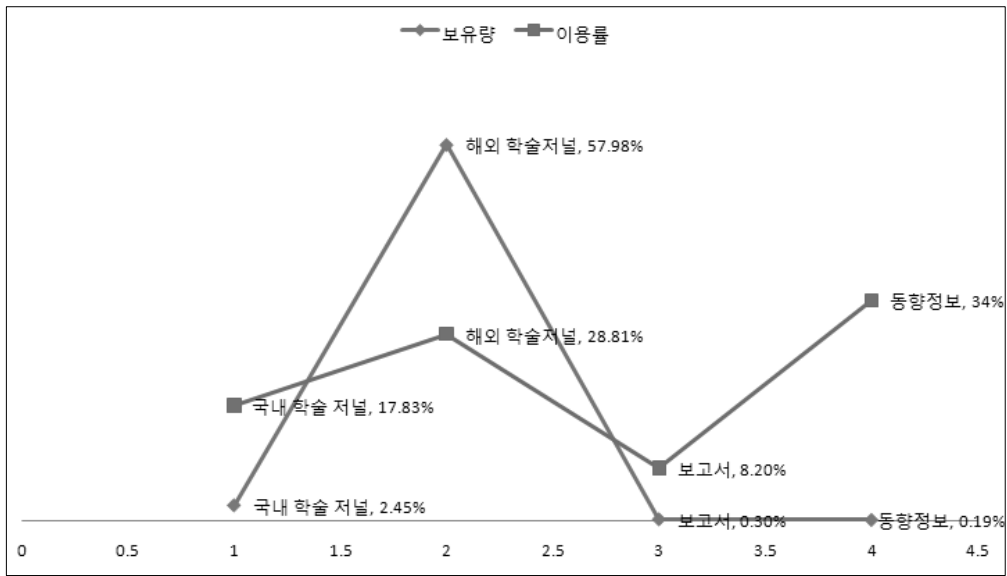
KISTI에서 보유하고 있는 자원들 중에서 대부분이 해외 학술저널들로 구성되어 있고 기타에서는 사실정보, 산업표준으로 구성 되어 있다. 2010년 1월 기준 KISTI가 보유하고 있는 전체 데이터 약 9천만 건에 대한 비율은 〈표 5〉와 같다. 해외 학술저널이 57.98%, 국내 학술저널이 2.45% 정도로 나타나고 있다. 국내 학술저널은 해외 학술저널에 비해 상당히 낮은 분포를 나타내고 있다. 해외/국내 학술저널에 비해서 연구보고서는 0.3%, 동향정보는 0.19%로 해외/국내 학술 저널에 비해 매우 낮은 분포를 나타내고 있다.

KISTI 정보 보유량과 사용자의 상세 이용 현황을 같이 분석하면 〈그림 1〉과 같다. 〈그림 1〉은 전체 데이터 보유량과 이용률을 중심으로 나타낸 것으로 해외 학술저널은 데이터 보유량에 비해서 이용률이 낮게 나타났고 동향정보, 연구보고서 및 국내 학술저널 등은 보유량에 비해 이용률이 높게 나타났다.

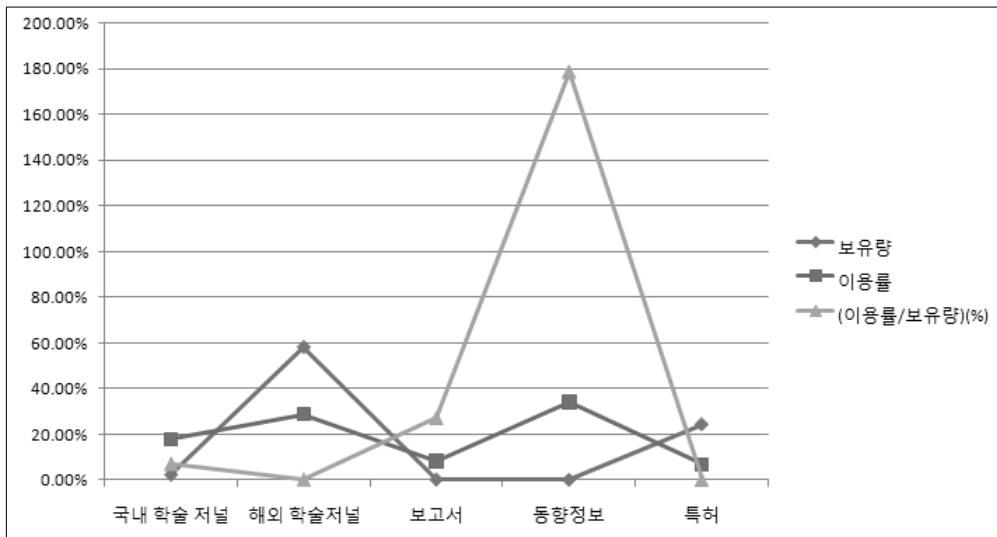
〈그림 1〉을 중심으로 전체 데이터 보유량을 중심으로 이용률을 나눈 결과는 〈그림 2〉와 같고 그 척도는 “(이용률/보유량)%”으로 분석이 되어 진다. 이와 같은 결과로 보면 동향정보(178.94%), 연구보고서(27.33%), 국내 학술 저널(7.27%), 해외 학술저널(0.49%)이고 특허(0.27%) 순으로 나타났다. 따라서 데이터 보유량과 이용률을 바탕으로 분석할 경우에 디지털 아카이빙 우선순위는 동향정보, 연구보고서, 국내 학술저널, 해외 학술저널 및 특허를 중심으로 아카이빙 대상을 선정할 수 있다.

〈표 5〉 전체 데이터 보유량 비율

구분	국내 학술 저널	해외 학술 저널	연구보고서	동향정보	특허	기타	전체 건 수
보유량	2,240,105	53,001,421	272,912	173,427	22,257,181	13,462,083	91,407,129
백분율(%)	2.45	57.98	0.03	0.19	24.35	0.14	100.00



〈그림 1〉 각 자원별 보유량과 이용률 분석 결과



〈그림 2〉 보유량을 기준으로 이용률 나눈 것

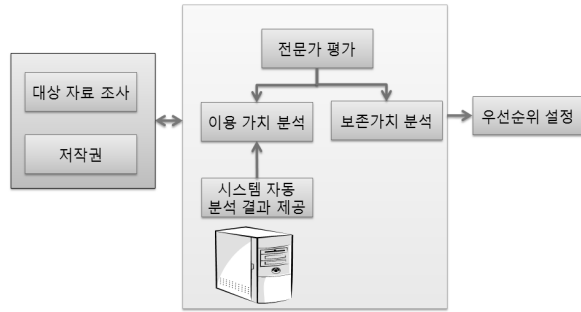
디지털 자원의 아카이빙을 위한 우선순위 결정에는 많은 요소들을 고려해야하지만 그 중에서도 이용가치와 보존가치, 저작권 등이 중요한 요소로 고려되어진다.

본 연구에서 자원의 이용가치를 중심으로 자원의 사용량이나 보유량 등의 정보는 시스템에서 자동으로 추출할 수 있는 요소들이다. 따라서 이와 같은 요소들을 시스템에 반영할 경우에 보다 전문가들이 평가하기에 용이하며 시스템에서 자동으로 제시함으로써 전문가가 아카이빙 우선순위 설정을 위한 평가 시스템에서 평가를 하기위해서 기초자료로 자동으로 분석하여 제공할

수가 있다.

디지털 자원의 아카이빙 평가 모델은 〈그림 3〉과 같다. 〈그림 3〉에서 제시한 것과 같이 우선 대상 자료를 조사와 함께 저작권 여부를 조사하여야 한다. 그리고 이용가치 분석과 보존가치 분석을 이용하여 전문가가 대상 자료에 대한 평가를 한다. 그리고 시스템에서 자동으로 제공 될 수 없는 것들에 대해서는 전문가 직접 평가를 한다. 예를 들면 경제적/산업적 파급효과 등과 같은 요소를 고려할 수가 있다. 전문가 평가에서 부적합으로 판정이 될 경우에는 대상 자료 재조사 또는 우선순위 설정

에서 제외가 된다. 이렇게 전문가들의 평가를 기반으로 우선순위가 설정이 되면 대상 자원의 아카이빙을 시행한다.



〈그림 3〉 디지털 아카이빙 자원 우선순위 평가 모델

5. 결론 및 제언

디지털 자원들은 기하급수적으로 늘어나고 있지만 국가적 차원에서 효율적으로 관리하지 않으면 급속한 폐기와 망실이 발생할 것이다. 미래세대를 위하여 디지털 자원의 생산뿐만 아니라 항구적 보존과 접근을 위한 더욱더 많은 관심과 노력이 요구되고 있다.

디지털 자원들을 아카이빙하기 위해서는 많은 확고한 아카이빙 정책과 전문인력, 지속적인 예산 지원, 기술개발 등이 뒷받침되어야 하고, 국가의 대표적인 자료수집 및 보존기관간의 협력이 중요하다. 또한 각 기관이나 단체에 적합한 아카이빙 대상 자원의 우선순위를 고려한 아카이빙시스템이 구축되어야 한다.

본 연구에서는 아카이빙 우선순위에 고려해야할 다양한 요소들 중에서 보존가치보다는 이용가치를 중심으로 분석을 하였다. 즉, 현재 이용가치가 높은 자원은 미래의 이용 잠재성이 높다는 전제하에서 KISTI 자원들 중에서 아카이빙 우선순위 결정을 위한 분석을 하였다. 단순하게 이용률만을 사용하지 않고 데이터 보유량에 비해서 얼마나 많이 이용되고 있는지를 중심으로 분석한 결과 동향정보, 연구보고서, 국내 학술저널, 해외 학술저널 및 특허 순으로 나타났다.

각 기관별로 디지털 자원의 아카이빙을 위한 대상 자원의 선정을 위해서는 다양 요소들이 고려되어야 하고 지속적인 연구가 필요하다. 실제적인 연구를 위해서는 주요 보존기관간의 협력과 실무 담당자의 면담 조사를 통하여 그 기관의 정책과 경제적, 산업적 파급효과 등을

함께 고려한 모델을 연구해야한다. 하지만 너무나 많은 요소들을 고려할 경우에는 그 복잡도가 높아지게 되어 시스템적으로 반영하기가 어렵다. 따라서 국가의 주요 기관간 협력을 통한 정책적 접근과 기관의 특성에 적합한 모델을 선정해야하는 것이 바람직하며, 이용가치 측면에서는 이용률과 보유량 등을 함께 고려할 필요성이 있다.

【참고 문헌】

김갑성. 2002. 『과학기술정보 수요조사 및 정보가치도 작성연구』. 서울: 한국과학기술정보연구원, 비매품.
네덜란드국립도서관 홈페이지. [cited 2010.7.11].
〈http://www.kb.nl/hrd/dd/index-en.html〉.
일본과학기술진흥기구 홈페이지. [cited 2010.7.10].
〈http://www.journalarchive.jst.go.jp〉.
정영임, 최호남, 최선희. 2009. 아카이빙 데이터의 활용성 증진을 위한 전략연구: 국내외 학술논문을 중심으로. 『정보관리학회지』, 27(1): 185-206.
조재인. 2004. 지식정보자원 보존체계에 관한 연구. 『한국도서관·정보학회지』, 35(4): 85-104.
차용진. 2001. 국가지식정보자원 디지털화 계획 수립방안에 관한 연구: 전략적 DB 구축분야를 중심으로. 『사회과학논총』, 11: 323-350.
최호남, 이응봉. 2005. 해외 전자저널의 디지털 아카이브 구축 전략에 관한 연구. 『한국문헌정보학회지』, 39(2): 161-183.
한국과학기술정보연구원. 2005. 『국가 디지털 아카이빙 체제 구축에 관한 연구』. 대전: 동연구원, 연구보고서.
한국과학기술정보연구원. 2010. 『수명주기 기반 디지털 콘텐츠 아카이빙 정책 연구』. 대전: 동연구원, 연구보고서.
Hodge, G. and E. Frangakis. 2004. Digital preservation and permanent access to scientific information: the state of the practice, 'ICSTI & CENDI.' Information Bridge Homepage [cited 2010.7.10].
〈http://www.osti.gov/bridge〉.
LOCKSS Homepage. [cited 2010.2.10].
〈http://www.lockss.org/lockss/Home〉.
Marcel Ras. 2008. The KB e-Depot: Building and Man-

- aging a Safe Place for e-Journals. *Liber Quarterly*, 19(1): 44-53.
- NASA Technical Reports Server(NTRS) Homepage. [cited 2010.7.11]. <<http://ntrs.nasa.gov>>.
- Portico Homepage. [cited 2010.7.10]. <<http://www.portico.org/digital-preservation>>.
- Technical Reports Archives and Image Library Homepage. [cited 2010.7.10]. <<http://trail.gwla.org>>.