대학도서관의 연구데이터관리서비스에 관한 연구
- 미국 연구중심대학도서관을 중심으로 -

A Study on Research Data Management Services of Research University Libraries in the U.S.

김 지 현 (Jihyun Kim)*

초 록

본 연구에서는 미국 연구중심대학도서관에서 최근 구축・운영되고 있는 연구데이터관리(Research Data Management, RDM) 서비스의 구성요소와 웹사이트에서 제공되는 서비스의 내용을 분석하여 그 현황을 조사하였다. 미국 카네기분류에서 제시된 최고연구중심대학(RU/VH) 중 RDM 서비스를 제공하는 31개 대학도서관의 서비스 내용을 분석하였으며, 분석의 기준은 선행연구에서 제시된 서비스의 구성요소 9가지로 (1) DMP 작성지원; (2) 데이터 파일 정리; (3) 데이터 기술; (4) 데이터 저장; (5) 데이터 공유 및 접근; (6) 데이터 보존; (7) 데이터 인용; (8) 데이터관리 교육; (9) 데이터 지적재산권을 포함한다. 분석 결과 대다수의 기관에서 DMP 작성지원서비스를 운영하고 있었으며 절반 이상의 기관에서 데이터 기술과 데이터 보존 및 데이터관리 교육 서비스를 제공하고 있었다. 특히 학문분야별 데이터에 적합한 메타데이터 및 레포지터리로의 안내와 워크숍 또는 개별 컨설팅을 통한 교육에 초점을 맞추고 있음을 확인할 수 있었다. 향후 기관 내 혹은 기관 간 협력에 기반을 둔 RDM 서비스 운영과 서비스 담당자의 지식 및 역량에 대한 논의와 연구가 지속되어야 할 것이다.

ABSTRACT

This study examined the current practices of Research Data Management (RDM) services recently built and implemented at research university libraries in the U.S. by analyzing the components of the services and the contents presented in their web sites. The study then analyzed the content of web pages describing the services provided by 31 Research Universities/Very High research activity determined based on the Carnegie Classification. The analysis was based on 9 components of the services suggested by previous studies, including (1) DMP support; (2) File organization; (3) Data description; (4) Data storage; (5) Data sharing and access; (6) Data preservation; (7) Data citation; (8) Data management training; (9) Intellectual property of data. As a result, the vast majority of the universities offered the service of DMP support. More than half of the universities provided the services for describing and preserving data, as well as data management training. Specifically, RDM services focused on offering the guidance to disciplinary metadata and repositories of relevance, or training via individual consulting services. More research and discussion is necessary to better understand an intra- or inter-institutional collaboration for implementing the services and knowledge and competency of librarians in charge of the services.

키워드: 연구데이터, 연구데이터관리서비스, 대학도서관
Research Data, Research Data Management Services, University Libraries

---

* 이화여자대학교 사회과학대학 문헌정보학 전공 조교수(kim.jh@ewha.ac.kr)
논문접수일자 : 2014년 9월 1일 논문심사일자 : 2014년 9월 1일 게재확정일자 : 2014년 9월 18일
한국비블리아학회지, 25(3): 165-189, 2014. [http://dx.doi.org/10.14699/kbiblia.2014.25.3.165]

www.kci.go.kr

1. 서 론

1.1 연구의 목적

연구자들의 연구 활동을 통해 생산되는 데이터는 학술지논문 등 최종연구결과물의 기반이 되는 원 자료이며 다른 학문적 관점을 가진 연구자에 의해 새로운 연구결과를 창출할 수 있다는 점에서 그 활용 가치가 매우 높다. 미국, 영국, 캐나다 등지에서는 이러한 연구데이터의 가치에 주목하여 2000년대 초반부터 연구데이터의 관리, 공유 및 보존과 관련된 연구와 실무에 대하여 국가적 차원의 지원을 제공하여 왔다. 미국의 경우 미국국립과학재단(National Science Foundation, 이하 NSF)과 미국보건기구(National Institute of Health, NIH) 등 주요 정부연구비지원기관에서 연구자들에게 논문뿐만 아니라 논문의 기초가 되는 데이터의 공유를 지속적으로 독려해왔다. 2010년 5월 NSF에서는 연구비계획서를 제출하는 모든 연구자들이 데이터관리계획(Data Management Plan, 이하 DMP)을 의무적으로 제출하게 하는 데이터 공유 정책 개정안을 발표하였다.

이러한 변화는 미국 연구기관 및 대학에 속한 도서관의 역할에 대한 논의로 이어졌으며 미국연구도서관협회(Association of Research Libraries, ARL)에서는 NSF의 DMP 의무사항을 준수하려는 연구자들에게 제공될 수 있는 서비스와 안내에 대한 가이드를 제공하였다. 이 가이드에서 NSF의 정책 변화는 데이터 중심적이고 협력 기반의 연구 환경 변화와 함께 연구결과물의 공개와 배포를 활성화하려는 목적으로 이루어졌음을 언급하고 있다(Hswe and Holt 2011).

이와 같이 연구자들의 DMP 작성을 지원하기 위한 서비스를 포함하여 최근 미국의 연구중심대학도서관에서는 대학 내 연구자들이 생산하는 데이터의 관리와 보존을 보다 적극적으로 지원하기 위한 연구데이터관리서비스(Research Data Management Service)를 구축・운영하는 사례가 증가하고 있다. 이는 미국에서 그간 데이터 관리와 보존 영역에서 수행된 수많은 연구와 실무 사례를 바탕으로 연구데이터관리의 모범실무(best practices)가 정립되고 있음을 보여준다. 따라서 본 연구에서는 이러한 서비스를 운영하고 있는 미국 연구중심대학 도서관 웹사이트에서 제공되는 서비스의 내용을 선행연구에서 제시된 연구데이터관리서비스 구성요소를 기준으로 분석하였다. 이를 통해 미국 연구중심대학 도서관에서 공통적으로 제시하고 있는 서비스의 내용을 확인하였으며 향후 서비스가 나아가야 할 방향을 논의하였다.

국내에서는 한국과학기술정보연구원(KISTI)에서 과학데이터 공유와 보존을 위한 시스템인 P-CUBE(Platform for Convergence research and Unification of Big E-resources)를 운영하고 있으며 한국연구재단의 기초학문자료센터(Korea Research Memory, KRM)도 인문사회분야 연구의 원 자료를 수집하는 등 정부기관을 중심으로 연구데이터관리서비스가 진행되고 있으나 아직 이러한 서비스가 본격적으로 활성화되지는 않은 실정이다. 국내 연구데이터관리서비스의 체계 및 운영 개선을 위해 본 연구의 분석결과가 기초자료로 활용될 수 있을 것이다.

www.kci.go.kr

1.2 연구의 범위와 방법

미국 내 대학을 구분할 때 근거자료로 삼는 대표적인 자료인 카네기 분류(Carnegie Classification)는 카네기재단에서 고안하였으며 교육조사 및 분석 활동에 주로 활용된다. 카네기 분류는 미국 내 대학을 학부교육, 대학원교육, 규모 등 특정 측면을 기준으로 분류한 6개의 분류체계와 1970년부터 카네기분류로서 활용되어 온 전통적인 분류체계를 수정・보완한 기본분류(Basic classification)를 포함한다(Carnegie Foundation 2014). 본 연구에서는 기본 분류에서 제시된 박사학위수여대학(Doctorate-granting university) 중에서 가장 활발히 연구를 수행하는 기관들을 분류한 최고연구중심대학(RU/VH: Research University-Very High research activity)인 108개 대학의 도서관 웹사이트를 살펴보았다. 이 대학들은 미국 내에서 박사학위 배출인원, 연구교수인원, 연구개발 투자 등을 기준으로 가장 활발히 연구를 수행하는 기관이다. 따라서 대학 내 연구자들의 연구 활동에서 생산되는 데이터의 관리서비스에 많은 관심을 가지고 이를 수행하고 있을 것으로 가정하여 본 연구의 조사대상으로 삼았다.

웹사이트 조사 결과 이들 108개 대학 중에서 연구데이터관리서비스에 대한 설명 또는 가이드를 제공하고 있는 대학도서관은 모두 31개 기관(29%)인 것으로 나타났다. 본 연구에서는 이들 31개 대학도서관 웹사이트에서 제공되는 연구데이터관리서비스의 내용을 선행연구에서 공통적으로 제시된 연구데이터관리서비스의 구성요소를 기준으로 분석하였다. 분석의 기준은 (1) DMP 작성지원; (2) 데이터 파일 정리; (3) 데이터 기술(description); (4) 데이터 저장; (5) 데이터 공유 및 접근; (6) 데이터 보존; (7) 데이터 인용; (8) 데이터관리 교육; (9) 데이터 지적재산권의 9가지 요소를 활용하였다.

2. 대학도서관의 연구데이터 관리서비스

2.1 연구데이터관리서비스의 정의 및 필요성

연구데이터관리(Records Data Management, 이하 RDM)서비스는 연구자들과 그들이 속한 기관에서 연구데이터를 생산, 수집, 이용, 보존하는데 있어 방향을 제시하고 지원하는 인력, 정책, 자원 및 기술을 포괄하는 시스템으로 정의할 수 있다(Steeleworthy 2014). 이는 최근 해외의 정부연구비기관에서 의무사항으로 제시하고 있는 데이터관리계획(Data Management Plan, 이하 DMP)의 작성 지원, 데이터 컬렉션과 레포지터리 구축, 디지털 보존과 데이터 리터러시를 포괄하는 광범위한 활동을 포함한다(Witt 2012). 데이터 저장이나 파일 보안 등 연구과제 수행 중에 발생하는 데이터관리 요구사항을 지원하며 데이터에 대한 메타데이터 제공과 데이터의 공유를 지원하는 기능도 수행한다(Fearon 외 2013).

이러한 RDM 서비스는 연구비지원기관에서 연구비지원자의 DMP의 제출을 의무사항으로 규정하면서 DMP 작성을 지원하기 위한 협의의 서비스로서 규정되기도 한다. 그러나 데이터 관리의 전 단계를 포괄하는 RDM 서비스를 통해 데이터가 과제수행기간 혹은 그 이후에도 검

www.kci.go.kr

색이나 접근이 가능한 형태로 장기간 관리될 수 있다. 또한 협력연구가 증가하면서 다양한 연구팀에서 운영되는 플랫폼과 시스템 간 데이터 교환의 필요성이 확대되면서 그 중요성은 더욱 강조되고 있는 실정이다(Jones 외 2013).

이와 더불어 RDM 서비스는 연구 및 학술환경의 변화를 적극적으로 반영하는 대학도서관 서비스로서 의의가 있다. Lyon(2012)은 가장 핵심이 되는 학술커뮤니케이션의 변화로서 데이터 중심 연구환경(Data centric research landscape)을 꼽았으며, 대학도서관이 연구데이터관리라는 난제를 효과적으로 해결하기 위해 어떠한 변화를 모색해야 하는지를 다음의 세 가지로 제안하였다.

첫째, 기관 내에서 일어나는 연구 활동의 생애주기를 파악하여 각 연구 단계에서 어떠한 데이터관리의 요구사항이 발생하는지를 조사하고 이러한 요구사항을 반영하는 RDM 서비스를 구축하여야 한다.

둘째, 데이터관리를 담당하는 직원은 특정 분야 연구 활동 및 과정과 기술적인 표준, 메타데이터, 용어에 대한 충분한 지식을 갖추어야 한다. 또한 해당 분야의 데이터 레포지터리 또는 데이터센터를 인지하고 분야 내 대표 학술지에서의 데이터 출판 의무사항을 이해하고 있어야 한다. Lyon은 현재 도서관에 근무하는 직원들이 제한적이나마 이러한 역할을 담당할 수는 있겠지만 향후 문헌정보학 분야 교육과정에서 데이터관리를 위한 기술 및 지식을 반영하는 것이 바람직하다고 보았다.

셋째, 대학도서관은 동료평가, 오픈액세스, 풍부한 콘텐츠, 데이터에 대한 접근제공 및 메타데이터 제공이라는 다섯 가지 요소를 제공함으로써 학술 커뮤니케이션의 핵심주자 역할을 할 수 있을 것이라 전망하였다. 특히 대학도서관에서 과학데이터와 학술논문과 관련하여 일반대중들도 이해할 수 있는 요약정보를 제공함으로써 서비스 대상을 일반대중으로까지 확대할 수 있는 기회가 될 수 있다고 주장하였다.

이와 같이 RDM 서비스는 대학도서관이 변화하는 연구 환경에 적응하여 발전하는 하나의 서비스 방안으로서 주목받고 있으며 효율적인 RDM 서비스 제공을 위해 문헌정보학 교과과정에서 데이터관리의 지식과 기술을 갖춘 인력양성에 대한 요구가 높아지고 있다. 또한 대학도서관이 RDM 서비스라는 새로운 영역과 함께 오픈액세스 및 메타데이터 제공 등 기존 의 서비스 영역까지 포괄하는 연구 활동 지원서비스를 제공함으로써 그 역할을 공고히 할 것을 요청받고 있다.

2.2 연구데이터관리서비스의 운영

미국과 영국, 캐나다의 대학도서관을 중심으로 RDM 서비스의 구축 사례가 증가하고 있으며 최근 서비스를 담당하는 사서들을 중심으로 이들 서비스의 운영과 관련된 사례연구가 진행되어왔다. 캘리포니아 대학도서관의 RDM 서비스 구축 사례에서는 데이터관리의 생애주기를 계획, 데이터 수집, 데이터 관리와 공유, 데이터 공개의 4단계로 제시하고 각 단계에서 필요한 도구와 서비스 및 기능을 제안하였다. 또한 캘리포니아 디지털 도서관(California Digital Library, CDL)과의 파트너십을 기반으로 하는 협력 프레임워크를 통해 성공적인 RDM 서비스 운영이 가능할 것이라 전망하였다(Starr 외 2012). 이

www.kci.go.kr

와 같이 RDM 서비스 구축을 위한 기관 내 부서 간 협력뿐만 아니라 기관 간 협력을 논의한 연구도 존재한다. Raboin 외(2012)는 두 대학도서관을 중심으로 하는 통합적인 RDM 서비스 모형을 제안하였다. 이를 통해 데이터관리 교육 제공의 범위를 확대할 수 있고 사서들도 RDM 서비스를 제공하기 위해 지식과 기술을 습득하는 등 경력 개발의 기회가 될 것으로 보았다. 그러나 이러한 기관 간 협력모형에서 행정적인 지원을 얻는 것과 여러 가지 서비스를 통합하는 문제, 그리고 변화하는 이용자의 요구에 지속적으로 대응해야 한다는 점을 어려움으로 지적하였다.

RDM을 위한 협력 및 통합서비스 체제 하에서 참여하는 직원이 속한 부서 간 문화의 차이가 서비스에 대한 관점의 차이로 이어질 수 있다. 이를 논의한 연구로 Verbaan과 Cox(2014)는 영국의 연구중심대학에서 RDM 서비스에 참여하는 직원 20명을 대상으로 인터뷰를 실시하였다. 이들 직원들은 도서관, IT 서비스, 연구처(research office)에 소속되어 있으며 각 직원이 소속된 부서에 따라 RDM 서비스를 바라보는 관점에 차이가 있었다. 도서관 직원들은 RDM 서비스를 오픈액세스와 정보 리터러시 서비스의 확장으로 보고 있었으며 이를 새로운 서비스 영역으로 규정해야 한다고 주장하였다. 이에 비해 다른 부서의 직원들은 RDM 서비스에 소요되는 인력과 자원 부족을 우려하고 있었으며 RDM 서비스를 위험부담이 큰 과제로 인식하고 있었다. RDM 서비스와 관련하여 도서관 직원들은 보존과 홍보를 강조하는 반면 IT 직원들은 단기간의 데이터 저장에 초점을 맞추고 있었으며 연구처 직원들은 연구비 지원기관의 요구사항 준수와 연구실적 평가의 측면을 강조한다는 점에서도 차이가 존재하였다. 효율적인 RDM 서비스를 구축하기 위해서는 여러 관련 부서 간의 상호작용, 업무 스타일 및 서비스담당자에게 필요한 지식과 기술에 대한 심도 있는 논의가 필요함을 제언하였다.

RDM 서비스 중에서 데이터관리 교육에 대한 연구로서 Adamick 외(2012)는 미국 매사추세츠 앰허스트 대학의 대학원생들을 대상으로 일반적인 데이터관리에 대한 워크숍을 제공하였는데 그 결과 일반적인 내용과 함께 특정 연구 분야의 데이터관리와 관련된 예시와 도구를 안내하는 워크숍에 대한 요구가 높았던 것으로 나타났다. 이와 유사하게 영국 JISC와 DCC의 지원을 받아 수행된 Data Management Skills Support Initiative(DaMSSI)에서는 박사후 연구자들의 위한 데이터관리 교육프로그램 개발에서 일반적인 데이터 관리 능력을 핵심 지식으로 하면서 특정 분야의 사례를 제시하여 참가자들의 관심을 환기시킬 것을 강조하였다(Molloy and Snow 2012). Johnston 외(2012)는 미네소타 대학에서 교수들을 대상으로 제공한 DMP 작성 지원 워크숍의 사례를 제시하면서 이러한 워크숍이 연구자 입장에서는 연구비 신청대상 자격을 유지할 수 있는 계속교육 요건으로서 유용하다고 설명하였다. 또한 교육을 제공하는 사서 입장에서는 경력 개발의 기회가 될 수 있음을 제시하였다.

2.3 연구데이터관리서비스의 구성요소

영국 데이터 큐레이션 센터(Data Curation Center, 이하 DCC)에서 개발한 RDM 서비스

www.kci.go.kr

개발 지침에서는 RDM 서비스의 구성요소를 〈그림 1〉과 같이 (1) RDM 정책과 전략(RDM policy and strategy); (2) 사업계획과 지속가능성(Business plan and sustainability); (3) 데이터 관리계획(DMP); (4) 현용데이터관리(Managing active data); (5) 데이터 선정 및 수집(Data selection and handover); (6) 데이터 레포지터리(Data repositories); (7) 데이터 기술(Data catalogues); (8) 지침, 교육 및 지원 제공(Guidance, training and support)의 8가지 요소로 제시하였다(Jones 외 2013). 〈그림 1〉의 상단에 표현되어 있듯이 서비스 구성요소 중 RDM 정책과 전략은 서비스의 운영과 지속가능성을 고려한 사업계획 수립에 영향을 미치게 된다. 또한 〈그림 1〉의 하단에 있는 안내, 교육 훈련 및 지원은 DMP, 현용데이터관리, 데이터 선정 및 수집, 데이터 레포지터리, 데이터 기술이 순환하면서 이루어지는 RDM 실무를 지지하는 인프라로 작용하게 된다.

제시된 각각의 RDM 서비스 구성요소와 관련된 업무는 〈표 1〉과 같이 정리할 수 있다. 첫째, 기관 내 여러 구성원들의 동의와 협조를 이끌어낼 수 있는 정책과 전략을 수립한다. 둘째, RDM 서비스에 소요되는 비용을 산정하고 비용 대비 효과 분석을 실시하며 서비스의 지속가능성을 고려한 장기 비용을 산출한다. 셋째, RDM과 관련된 안내, 지침 및 헬프데스크와 다양한 교육 프로그램과 컨설팅 서비스를 운영한

〈그림 1〉 RDM 서비스의 구성요소

*출처: Jones, S., Pryor, G. and Whyte, A. (2013). How to Develop Research Data Management Services - a guide for HEIs. *DCC How-to Guides*. Edinburgh: Digital Curation Centre. p.5 [cited 2014.9.14]. 〈http://www.dcc.ac.uk/resources/how-guides〉.

www.kci.go.kr

〈표 1〉 RDM 서비스 구성요소와 업무

서비스 구성요소	업무
RDM 정책 및 전략	전략개발
	정책개발
	데이터 정책과 전략의 조화 및 실행
사업계획 및 지속가능성	비용 산정 및 비용효과 분석 실시
	서비스 지속가능성과 이와 관련된 장기 비용 고려
안내, 교육 훈련 및 지원	안내 및 헬프데스크 운영
	교육 프로그램 개발
	컨설팅 서비스 운영
데이터관리계획(DMP)	DMP 요구 조사
	DMP 템플릿 및 안내 제공
	DMP 작성지원 도구 제공
	교육 및 컨설팅
현용데이터 관리	데이터 저장
	클라우드 저장 서비스
	웹하드 서비스 지원
	자체적인 RDM 플랫폼 구축
데이터 선정 및 수집	연구자들의 데이터 제공 독려
	홍보 및 안내 제공
	데이터 기탁 동의서 작성
데이터 레포지터리	기관 레포지터리 구축 및 운영
	외부 레포지터리 안내
데이터 기술	메타데이터 선정 및 적용

*출처: Jones, S., Pryor, G. and Whyte, A. (2013). 'How to Develop Research Data Management Services - a guide for HEIs'. *DCC How-to Guides*. Edinburgh: Digital Curation Centre. pp.6-20. [cited 2014.9.14]. 〈http://www.dcc.ac.uk/resources/how-guides〉.

다. 넷째, DMP 작성과 관련된 연구자들의 요구를 조사하고 템플릿과 안내 및 DMP 작성을 지원하는 도구를 제공하며, 이와 관련된 교육과 컨설팅 서비스를 수행한다. 다섯째, 현용데이터 관리는 연구자들이 현재 연구 활동에서 이용하고 있는 데이터의 저장과 관리를 지원하는 기술적인 인프라로서 클라우드 서비스와 웹하드 서비스, 자체적인 RDM 플랫폼 구축을 제시하고 있다. 여섯째, 데이터 선정과 수집을 위한 기준을 마련하고 연구자들의 데이터 제공을 독려하며 이에 대한 홍보와 안내, 연구자들의 참여를 돕는 도구와 기탁(deposit)과 관련된 조건사항을 명시한 기탁동의서를 활용한다. 일곱째, 데이터의 보존 및 공유를 위한 기관 레포지터리를 구축, 운영하며 적절한 외부 레포지터리 서비스로 안내한다. 여덟째, 적합한 메타데이터의 선정과 이를 부여하는 워크플로우 개발 및 검색지원 서비스를 제공한다.

또 다른 연구로 Lyon(2012)은 연구 활동 생애주기에 기반을 둔 RDM 서비스의 구성요소를 10가지로 제시하였고 각각의 정의를 〈표 2〉와 같이 제안하였다. 이러한 서비스를 정착시키

www.kci.go.kr

기 위해서는 기관 내 여러 부서와의 협력관계를 조성하고 공통의 이해를 바탕으로 일관된 서비스를 제공하는 것이 중요하다. 또한 유사한 서비스 간의 통합적인 운영을 유도하고 RDM 계획 자료나 교육 자료를 공유함으로써 RDM에 대한 기관 구성원들의 학습효과를 높이는 것이 필요하다.

이와 더불어 실제 제공되는 RDM 서비스의 현황과 관련하여 미국연구도서관협회(Association of Research Libraries, ARL)에서는 ARL 회원도서관을 대상으로 RDM 서비스에 대한 설문조사를 실시하였다. 125개 회원도서관 중 73개 기관의 설문응답을 분석한 결과 54개 기관에서 RDM 서비스를 제공하고 있는 것으로 나타났다. 제공하는 RDM 서비스의 유형은 크게 (1) 교내 연구자들의 DMP 작성을 지원하기 위한 온라인 자료 제공, 컨설팅 및 교육 프로그램 제공; (2) DMP 작성 지원 이외의 서비스; (3) 기관 레포지터리에서의 데이터 보존 서비스; (4) 데이터 보존에 특화된 레포지터리 제공으로 구분하여 조사하였다. 이 중에서 데이터 보존에 특화된 레포지터리 운영을 제외하고는 절반 이상의 응답기관에서 상기한 서비스를 제시하고 있다고 응답하였다.

특히 DMP 작성 지원 이외의 서비스는 매우 다양한 서비스 영역을 포함하는데 (1) 연구자의 데이터에 적합한 메타데이터 선정 및 적용 지원; (2) 파일포맷 결정, 파일명 부여와 정리 지원; (3) 데이터 인용; (4) 데이터 공유와 접근; (5) 데이터 저장 및 백업; (6) 데이터관리 모범실무(best practices)에 대한 교육; (7) 데이터관리 모범실무에 대한 온라인 자료 제공; (8) 데이터 보안 및 익명화(anonymizing) (9) 데이터 권한 관리(data right management) 서비스를 수행하고 있는 것으로 나타났다(Fearon et al. 2013).

또 다른 설문조사로서 Cox와 Pinfield(2013)은 영국의 81개 대학도서관에서 응답한 설문 내용을 바탕으로 RDM 서비스의 현황을 조사하였다. 이들 대학도서관에서 제공하는 RDM 서비

〈표 2〉 RDM 서비스 요소와 정의

요소	정의
RDM 요구사항(requirements)	데이터 감사 및 평가도구를 활용하여 기관 내의 각 부서에서 생산되는 연구데이터 관리에 대한 요구사항을 조사
데이터관리계획(planning)	기관 내 연구자들이 효과적인 데이터관리계획을 수립할 수 있도록 지원과 지침 제공
데이터 기술 지원(informatics)	특정 학문분야에 적합한 메타데이터 표준, 데이터 포맷과 온톨로지 등 데이터 기술(description)을 위한 전문성 제공
데이터 인용(citation)	DataCite와 같이 데이터세트에 대한 영구적인 ID를 제공하여 데이터의 검색, 인용 및 재이용을 촉진시키는 제 3자 서비스에 대한 안내 제공
RDM 교육 훈련(training)	데이터관리 교육 프로그램 및 모듈과 교육 자료를 제공
RDM 라이선싱(licensing)	데이터 라이선싱에 대한 지침을 마련하여 데이터와 관련된 법적, 윤리적 문제 논의
데이터 선정 평가(appraisal)	데이터 저장 및 보존을 위한 선정 및 평가 작업에 대한 안내 제공
데이터 저장(storage)	기관 내 IT 서비스와 협력하여 데이터 저장장소에 대한 안내와 관련 인프라 제공
데이터 접근제공(access)	합법적인 범위 내에서 데이터 접근이 가능하도록 데이터 배포 및 엠바고 정책 제공
RDM 효과(impact)	데이터 재이용과 직접적으로 연관되는 RDM 서비스의 효과를 조사

www.kci.go.kr

스의 유형으로는 (1) 웹 포털 서비스; (2) 데이터 관리에 관한 컨설팅 서비스(advisory service); (3) RDM 계획 서비스; (4) 오픈 액세스 및 관련 정책 안내; (5) 데이터 인용; (6) 메타데이터; (7) 레포지터리; (8) RDM 감사(audit); (9) 재이용될 수 있는 데이터 안내(awareness of reusable sources); (10) 외부 데이터 소스 안내(external data sources); (11) 저작권; (12) 라이선싱; (13) 데이터 분석; (14) 데이터관리 교육; (15) 홍보를 통한 RDM 서비스 인식 제고; (16) 데이터 효과(data impact) 측정으로 제시되었다. 이 중에서 우선순위가 높은 순서대로 제시된 상위 5개의 서비스 유형은 오픈 액세스 및 관련 정책 안내, 저작권, 데이터 인용, 재이용될 수 있는 데이터 안내, 외부 데이터 소스 안내인 것으로 나타났다.

이들 네 편의 선행연구에서 나타난 RDM 서비스 구성요소를 비교하여 정리하면 〈표 3〉과

〈표 3〉 RDM 서비스 구성요소 비교

구성요소 \ 선행연구	영국 DCC RDM 서비스 개발지침 (Jones 외 2013)	Lyon (2012)	미국 ARL RDM 서비스 조사 연구 (Fearon 외 2013)	Cox and Pinfield (2013)
DMP 작성지원	• 데이터관리계획(DMP)	• RDM 계획	• DMP 작성지원	• RDM 계획 서비스
데이터 파일 정리 (organization)	—	• 데이터 기술 지원	• 파일포맷 결정, 파일명 부여와 정리	—
데이터 기술 (description)	• 데이터 기술(catalogues)	• 데이터 기술 지원	• 메타데이터 선정 및 적용	• 메타데이터
데이터 저장	• 현용데이터 관리	• 데이터 저장	• 데이터 저장 및 백업	—
데이터 공유 및 접근	• 데이터 선정 및 수집	• 데이터 접근제공	• 데이터 공유와 접근	• 오픈액세스 및 관련 정책 안내 • 재이용될 수 있는 데이터 안내
데이터 보존	• 데이터 선정 및 수집 • 데이터 레포지터리	• 데이터 선정 평가	• 기관 레포지터리 또는 데이터 레포지터리	• 레포지터리 • RDM 감사(audit)
데이터 인용	—	• 데이터 인용	• 데이터 인용	• 데이터 인용
데이터관리 교육	• 안내, 교육 훈련 및 지원	• RDM 교육훈련	• 데이터관리 모범실무에 대한 교육	• 컨설팅 서비스 • 데이터관리 교육
데이터 지적재산권	—	• RDM 라이선싱	• 데이터 권한 관리	• 저작권 • 라이선싱
데이터 재이용의 효과 측정	—	• RDM 효과	—	• 데이터 효과(data impact)
데이터 정책 및 전략 개발	• 데이터 정책 및 전략	—	—	—
사업계획 작성	• 사업계획 및 지속가능성	—	—	—
RDM 요구사항 조사	—	• RDM 요구사항	—	—
데이터 분석	—	—	—	• 데이터 분석
웹 포털 서비스	—	—	—	• 웹 포털 서비스
RDM 서비스 홍보	—	—	—	• 홍보를 통한 인식제고
보안 및 익명화	—	—	• 보안 및 익명화	—

www.kci.go.kr

같다. 비교 과정에서 하나의 서비스 구성요소가 두 가지 측면의 내용을 담고 있을 경우 중복하여 기재하였다. 예를 들어 영국 DCC RDM 서비스 개발 지침의 '데이터 선정 및 수집'은 연구자들의 데이터 제공을 독려하는 공유의 측면과 보존을 위해 데이터를 선정하는 업무를 포괄하고 있으므로 '데이터 공유 및 접근'과 '데이터 보존' 둘 다에 해당한다고 보았다. Lyon(2012)이 제시한 '데이터 기술 지원'의 경우에도 파일 포맷을 결정하는 데이터 파일 정리의 내용을 포함하고 있으므로 '데이터 기술'과 '데이터 파일 정리'에 중복 기재하였다.

〈표 3〉에서 제시된 바와 같이 두 편 이상의 연구에서 제시된 구성요소들은 모두 10개로 (1) DMP 작성지원; (2) 데이터 파일 정리; (3) 데이터 기술; (4) 데이터 저장; (5) 데이터 공유 및 접근; (6) 데이터 보존; (7) 데이터 인용; (8) 데이터관리 교육; (9) 데이터 지적재산권; (10) 데이터 재이용의 효과 측정이다. 이 중에서 데이터 재이용의 효과 측정은 서비스의 평가와 관련된 사항이며 아직 RDM 서비스가 초기 단계임을 감안하여 분석 기준에서 제외하였다. 따라서 본 연구에서는 네 편의 선행연구에서 공통적으로 제시된 서비스 구성요소 중 9개의 요소를 기준으로 미국 연구중심대학도서관 웹사이트에 제시된 RDM 서비스의 내용을 분석하고자 한다.

3. 대학도서관 연구데이터관리 서비스 내용 분석

본 절에서는 미국 카네기분류의 기본 분류체계에서 제시된 미국 내 박사학위수여대학 중 최고연구중심대학(RU/VH)으로 분류되어 있는 108개 기관 중에서 RDM 서비스에 대한 안내를 웹사이트에서 제시하고 있는 31개 기관을 중심으로 제시된 서비스의 내용을 분석하였다(〈표 5〉 참조). 분석의 기준으로 〈표 4〉에 제시되어 있는 RDM 서비스의 9개 요소를 활용하였다.

3.1 DMP 작성지원

미국의 주요 정부연구비지원기관에서는 연구계획서와 함께 DMP를 의무적으로 제출하도록 되어 있으며, 이는 연구비 지원자들이 연구

〈표 4〉 분석기준으로 활용한 RDM 서비스의 9개 요소

RDM 서비스 요소	정의
DMP 작성지원	연구자들의 DMP 작성을 보조, 지원하는 서비스
데이터 파일 정리	데이터 파일명 부여, 버전관리 및 파일포맷 선정 지원 서비스
데이터 기술	데이터에 적합한 메타데이터 표준 안내 서비스
데이터 저장	데이터의 저장 및 백업 관련 서비스
데이터 공유 및 접근	데이터의 공유 필요성과 공유방법에 대한 안내 서비스
데이터 보존	데이터 보존을 위한 레포지터리 안내 서비스
데이터 인용	데이터 인용 방법과 DOI 안내 서비스
데이터 관리 교육	데이터관리를 위한 개별 컨설팅, 워크숍 및 교육자료 제공 서비스
데이터 지적재산권	데이터와 관련된 저작권 및 라이선싱 안내 서비스

www.kci.go.kr

〈표 5〉 조사대상기관의 RDM 서비스 구성요소 분석

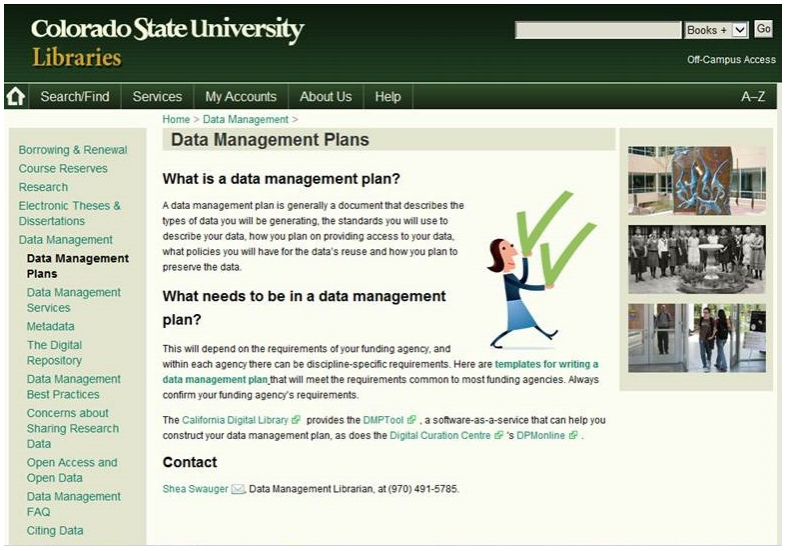
서비스 구성요소 / 대학	DMP 작성 지원	데이터 정리	데이터 기술	데이터 저장	데이터 공유와 접근	데이터 보존	데이터 인용	데이터 관리 교육	데이터 지적 재산권
카네기멜론 (Carnegie Mellon Univ.)	●		●	●					●
콜로라도주립 (Colorado State Univ.)	●	●			●		●	●	
코넬(Cornell Univ.)	●	●	●	●	●	●	●	●	●
듀크(Duke Univ.)	●		●		●		●	●	
플로리다주립 (Florida State Univ.)	●	●	●		●	●	●		●
하버드(Harvard Univ.)	●	●		●	●	●	●	●	
인디애나 블루밍턴 (Indiana Univ._Bloomington)				●	●	●		●	
메사추세츠공과대(MIT)	●	●	●	●		●		●	●
뉴욕(New York Univ.)	●		●	●				●	
펜실베이나주립 (Penn State Univ.)	●	●	●		●	●			
프린스턴(Princeton Univ.)	●							●	
퍼듀(Purdue Univ.)	●		●				●	●	
럿거스(Rutgers Univ.)	●		●		●	●	●	●	●
스탠포드(Stanford Univ.)	●	●	●	●	●	●		●	●
테네시(Univ. of Tennessee)	●	●	●	●	●	●	●	●	●
터프츠(Tufts Univ.)	●		●			●			●
뉴욕주립 버펄로 (Univ. at Buffalo)	●					●			
캘리포니아 로스앤젤레스 (UCLA)	●			●				●	
캘리포니아 샌디에고 (UC San Diego)	●	●	●		●	●	●	●	
신시내티(Univ. of Cincinnati)	●	●	●	●					
조지아(Univ. of Georgia)	●								
하와이(Univ. of Hawaii)									
휴스턴(Univ. of Houston)	●		●			●		●	
매사추세츠 앰허스트 (Univ. of Massachusetts_Amherst)	●	●	●				●	●	
마이애미(Univ. of Miami)	●		●			●	●		
미네소타(Univ. of Minnesota)	●	●	●			●	●	●	●
네브래스카 (Univ. of Nebraska)	●							●	
유타(Univ. of Utah)	●						●		
워싱턴(Univ. of Washington)	●					●	●		
위스콘신 매디슨 (Univ. of Wisconsin_Madison)	●	●	●	●	●	●	●	●	
버지니아 커먼웰스 (Virginia Commonwealth Univ.)	●					●		●	
기관 수 합계	29	13	19	11	12	18	15	20	9

www.kci.go.kr

계획서를 제출하는 단계에서부터 연구결과물의 생성, 관리 및 공유를 계획하게 하려는 취지에서 시행되고 있다. 이러한 추세를 반영하듯 31개 분석대상기관 중 거의 대다수인 29개 기관(93.5%)에서 DMP 작성지원을 위한 서비스를 제공하고 있었다. 서비스의 내용은 크게 (1) DMP의 정의와 필요성에 대한 설명; (2) DMP 작성을 위한 온라인 도구 안내; (3) 정부연구비 지원기관 또는 RDM 서비스 기관에서 제공하는 DMP 작성가이드 또는 예시로 나누어볼 수 있다.

DMP의 정의 또는 필요성을 설명하는 기관은 모두 9개 기관이었으며 콜로라도 주립대학 도서관의 경우 DMP를 제안하는 연구 활동에서 생산될 데이터의 유형, 데이터 기술을 위한 표준, 데이터 접근제공에 대한 계획, 데이터 재이용과 관련된 방침 및 데이터 보존 계획을 기술하는 문서라고 정의하고 있다(〈그림 2〉 참조). 캘리포니아 샌디에고 대학도서관에서는 DMP의 필요성으로 (1) 연구비지원기관의 요구사항 준수; (2) 연구의 영향력과 가시성 증가; (3) 필요한 데이터의 검색이 용이해지므로 연구 프로세스 간소화; (4) 데이터의 중복 감소; (5) 연구에 집중할 수 있는 시간 증가를 들었다.

DMP작성을 위한 온라인 도구는 29개 기관 중 23개 기관에서 제공하고 있었으며 23개 기관 모두 DMPTool에 대한 설명과 링크를 제공하고 있었다. DMPTool은 2011년 미국 캘리포니아, 일리노이-어바나샴페인, 버지니아 대학도서관과 스미스소니언, DataONE 프로젝트 및 영국 DCC의 파트너십에 기반을 두어 개발된 웹 기반 DMP 작성지원 도구이다. 미국 내 다양한 정부연구비지원기관의 DMP 요구사항에 맞는 DMP를 손쉽게 작성할 수 있으므로 개발 이후



〈그림 2〉 콜로라도 주립 대학도서관 DMP 안내 페이지

*출처: 〈http://lib.colostate.edu/services/data-management/data-management-plans〉. [cited 2014.9.14].

www.kci.go.kr

빠른 속도로 확산되어 활발히 이용되고 있다(Sallans and Donnelly 2012). 23개 기관 중 2개 기관은 영국 DCC에서 개발한 DMPOnline에 대한 링크도 함께 제공하고 있었다.

주요 연구비지원기관의 DMP 작성가이드 혹은 예시를 제공하는 기관은 19기관인 것으로 나타났으며 이 중에서 대다수를 차지하는 14개 기관에서 미국국립과학재단(National Science Foundation, NSF)의 가이드를 제공하고 있었다. 10개 기관에서는 미국보건기구(National Institute of Health, NIH)의 가이드와 예시를, 그리고 7개 기관에서는 사회과학분야 데이터아카이브를 제공하는 ICPSR(Interuniversity Consortium for Political and Social Research)의 DMP 템플릿을 제공하고 있었다. 14개 기관 중 10개 기관은 NSF, NIH 또는 ICPSR 가이드 중 두 개 이상을 함께 제시하고 있었다. 이 외에도 소수의 기관에서 미국국립인문재단(National Endowment for the Humanities, NEH), 미국해양대기청(National Oceanic and Atmospheric Administration, NOAA), 미국항공우주국(NASA) 등 DMP 제출을 의무사항으로 규정하고 있는 미국 내 다양한 연구비지원기관의 관련 정책을 안내하고 있었다.

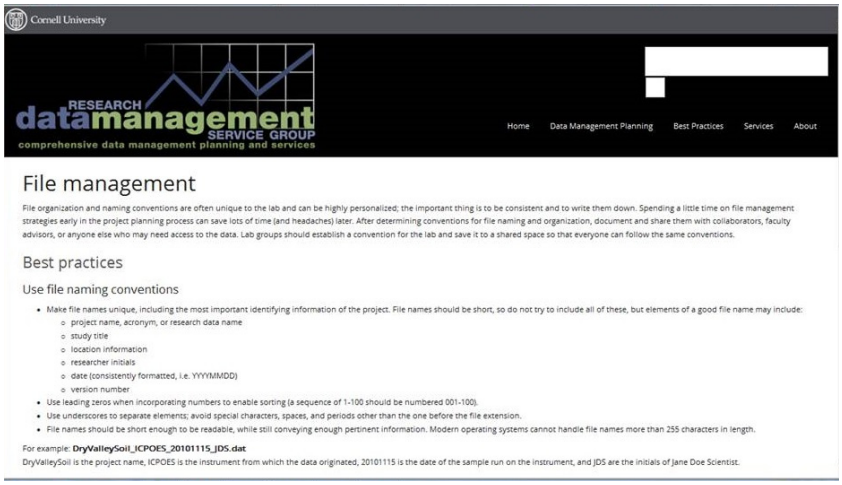
뿐만 아니라 MIT 대학도서관을 포함한 8개 기관에서는 MIT에서 자체적으로 개발한 DMP 체크리스트를 안내하고 있었다. 이 체크리스트에는 (1) 데이터의 유형; (2) 데이터의 양과 증가율; (3) 현재 또는 잠재적 데이터 이용자; (4) 데이터의 통제 주체; (5) 데이터의 유지 기한; (6) 데이터의 생산 및 분석에 사용되는 도구와 소프트웨어; (7) 연구비지원기관 또는 학술지에서 요구하는 데이터 공유 및 관리계획; (8) 파일명 부여 및 파일 포맷 결정 등 데이터 파일 정리 작업; (9) 데이터 저장 및 백업; (10) 데이터 공유와 관련된 이슈 또는 공유 방법을 담고 있다. 이는 연구자들이 연구비지원기관의 의무사항으로서만 DMP를 바라보는 것이 아니라 스스로 데이터관리에 있어 필요한 부분을 평가할 수 있다는 점에서 유용한 자료라고 할 수 있다.

3.2 데이터 파일 정리

데이터의 관리에 있어 데이터의 공유 및 보존에 적합한 파일포맷을 결정하고 일관된 파일명 부여 및 버전 관리 방식을 사용하는 등 데이터 파일을 조직하고 정리하는 작업이 필요하다. Ferguson(2012)은 분자생물학 분야 연구실 실험기구 내에 저장되어 있는 파일을 조사한 결과 독점적이고 상업적인 파일포맷으로 된 데이터가 절반 이상이었으며 파일명 부여의 방식도 불규칙인 것으로 나타났다. 데이터 파일 정리에 대한 안내와 교육이 RDM 서비스를 통해 제공될 필요가 있음을 알 수 있다.

데이터 파일 정리와 관한 내용을 제공하는 기관은 31개 기관 중 13개 기관(42%)인 것으로 나타났다. 데이터 파일 정리와 관련해서 거의 대다수인 12개 기관은 파일명 부여와 관련된 안내를 제공하고 있었으며 8개 기관은 파일포맷에 관련된 사항을, 5개 기관은 파일 버전 관리와 관련된 내용을 제시하고 있었다. 세 가지 모두를 다루고 있는 기관은 3개 기관이었으며 파일명 부여와 파일포맷에 대한 내용을 제시하는 기관은 6개 기관, 파일명 부여와 버전관리를 다루는 기관은 2개 기관이었다. 파일명 부여에 대한 내

www.kci.go.kr



〈그림 3〉 코넬 대학도서관 데이터 파일명 부여 안내 페이지

*출처: 〈http://data.research.cornell.edu/content/file-management〉. [cited 2014.9.14].

용만을 제시하는 기관은 2개 기관, 파일포맷에 대한 내용만 제시한 기관은 1개 기관인 것으로 나타났다.

파일명을 부여하는 방법으로 코넬 대학도서관에서 제시하는 내용을 살펴보면 짧고 특징적인 이름으로 하되 프로젝트명이나 약어, 논문제목, 저장장소 관련 정보, 연구자 이니셜, 날짜, 버전 정보 등을 포함하도록 하고 있다. 또한 일련번호는 정렬을 위해 0으로 시작하는 번호를 사용하고 이름의 요소를 분리할 때는 밑줄문자(underscore)를 사용할 것을 권장하고 있다(〈그림 3〉 참조). 버전관리와 관련하여 위스콘신 매디슨 대학도서관에서는 파일을 저장할 때마다 파일이름 끝에 v.01 형식의 버전 번호를 넣을 것과 최종 버전에는 FINAL을 번호대신 수록하고 자동적으로 파일 버전을 추적할 수 있는 버전관리 소프트웨어를 사용할 것을 제안하였다.

파일포맷의 경우 스탠포드 대학도서관에서는 공개되고 비독점적인(open and non-proprietary) 포맷을 사용할 것을 권장하고 있으며, 공개된 포맷으로 비압축(uncompressed), 비암호화(unencrypted)되고 연구 커뮤니티에서 일반적으로 활용되고 있는 것을 이상적인 데이터 파일포맷으로 설명하고 있었다. 또한 텍스트, 표 형태의 데이터(tabular data), 사진, 오디오, 동영상 등 데이터의 유형에 따라 선호되는 파일포맷의 예를 제시하였다.

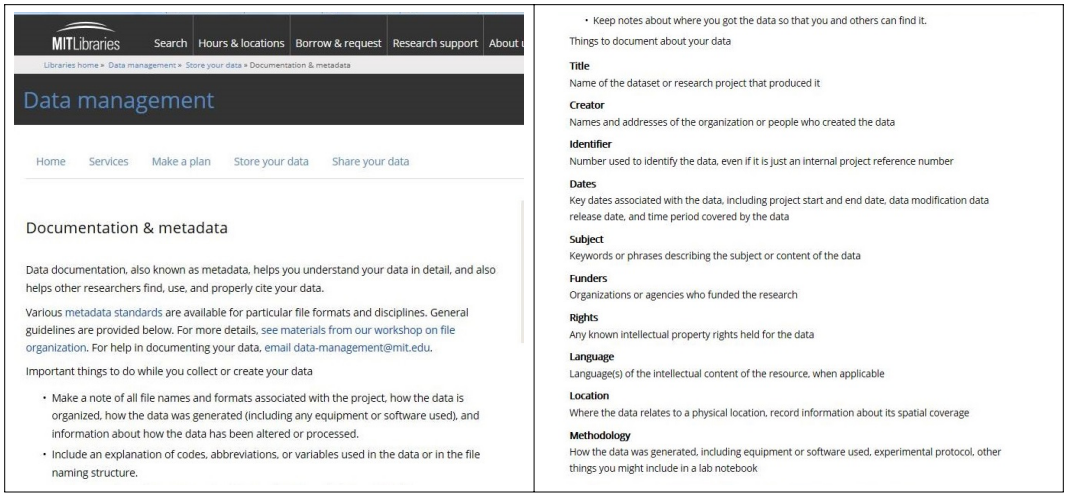
3.3 데이터 기술

데이터의 효율적인 검색과 접근을 위해 적절한 메타데이터 표준을 선정하고 이를 부여하는 작업은 데이터 관리에 있어 매우 중요하다고 할 수 있다. Parsons(2013)는 RDM 서비스를 운영할 때 DMP, 데이터 저장과 함께 데이터 기술 및 메타데이터에 대한 교육이 반드시 이루어져야 함을 강조하고 있다.

www.kci.go.kr

조사대상기관 중 19개 기관(61.3%)에서 데이터 기술과 관련된 내용을 제시하고 있었다. 우선 메타데이터의 정의를 제공하는 기관은 8개 기관이었으며 6개 기관에서 그 필요성을 언급하고 있었다. 예를 들어 카네기멜론 대학도서관에 따르면 메타데이터는 데이터에 대한 기술과 관련된 사람 또는 기관의 연락정보 및 데이터 이용조건을 포함하며 이러한 정보를 공식화하고 이를 데이터에 연결시키는 기능을 수행한다. 또한 연구비지원기관에서도 메타데이터 제공에 대한 사항을 DMP에 수록하도록 되어 있어 데이터에 대한 최소한의 메타데이터를 제공하는 작업이 반드시 필요함을 언급하고 있다. 데이터를 위한 일반적인 메타데이터 요소를 제시하는 기관은 5개 기관이었는데 MIT 대학도서관의 경우 메타데이터 요소를 제목, 생산자, 식별자, 날짜, 주제, 연구비지원기관, 권한, 장소, 방법론으로 제시하고 있다(〈그림 4〉 참조).

다양한 메타데이터 표준에 대한 설명도 제시되었는데 가장 자주 언급되었던 표준은 일반적으로 활용될 수 있는 더블린코어(Dublin Core)로서 모두 11기관에서 이를 제시하였다. 또한 10기관에서 사회과학분야에서 활용되는 DDI(Data Documentation Initiative)를 언급하였다. 생태학 분야에서 활용되는 EML(Ecological Metadata Language)와 생명과학 분야 특히 생명다양성(biodiversity) 분야에서 많이 활용되는 다윈 코어(Darwin Core)를 언급한 기관은 각각 6개 기관인 것으로 나타났다. 예술 및 인문학 분야의 메타데이터로 TEI(Text Encoding Initiative)와 VRC(Visual Resources Association Core)도 각각 3개 기관에서 언급되었다. 이 외에 도서관 분야 표준인 MODS(Metadata Object Description Schema)와 METS(Metadata Encoding and Transmission Standard)를 제시한 기관도 3곳인 것으로 나타났다.



〈그림 4〉 MIT 대학도서관 데이터 기술 안내 페이지

*출처: 〈http://libraries.mit.edu/data-management/store/documentation/〉. [cited 2014.9.14].

www.kci.go.kr

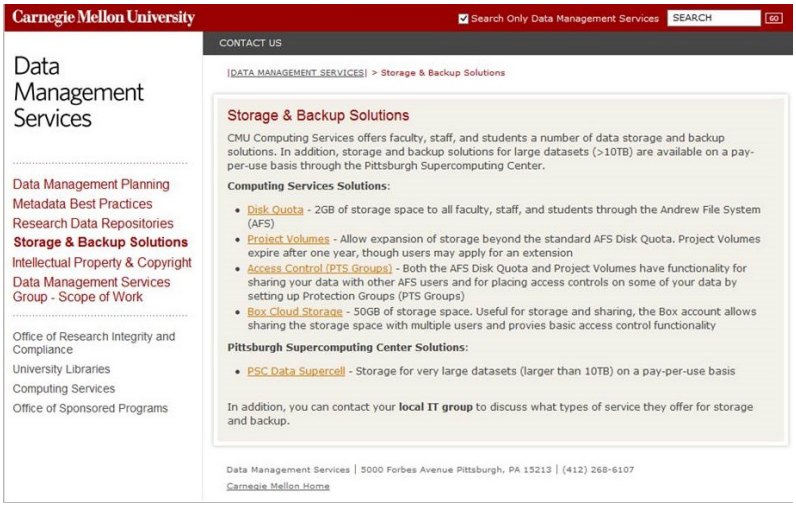
3.4 데이터 저장

데이터 저장 및 백업과 관련된 사항은 연구자들이 데이터를 관리하면서 겪는 어려움으로 자주 논의되며 RDM 서비스를 통해 이에 대한 교육과 적절한 자원으로의 안내가 필요하다. Jones 외(2008)는 영국 에든버러 대학 연구자들의 데이터관리 사례를 통해 연구자들이 데이터에 대한 백업을 제대로 수행하지 않아 데이터가 유실되거나 검색되지 않는 문제가 빈번하게 발생함을 지적하였다.

데이터 저장을 RDM 서비스 영역에 포함하는 기관은 모두 11기관(35.5%)인 것으로 나타났으며 이 중에서 5개 기관이 저장과 백업을 지원해주는 교내 관련 기관으로의 링크를 제공하고 있었다. 카네기멜론 대학도서관의 경우 데이터의 용량에 제한을 두어 2GB의 저장용량은 모든 대학구성원에게 제공되며 50GB까지의 클라우드 저장 서비스도 제공하고 있다. 10TB 이상의 대용량데이터의 경우 유료로 사용할 수 있는 서비스로 안내를 제공하고 있었다(<그림 5> 참조). 또한 3개 기관에서 백업에 대한 가이드를 제공하고 있었는데 데이터에 대한 3개의 복사본을 만들어 데이터가 원래 생산되었던 장소에 하나의 사본을 저장하고 나머지는 외부의 두 장소에 분산해서 보관할 것을 공통적으로 제시하고 있었다. 이 외에도 저장 또는 백업을 위한 솔루션을 제시하는 기관도 각각 2개 기관인 것으로 나타났다.

3.5 데이터 공유와 접근

데이터 공유와 접근을 통해 데이터의 재이용이 활성화되면 데이터관리의 궁극적인 목적인 연구영향력의 강화와 학문의 발전에 기여할 수 있다. 또한 연구자들은 데이터 공유를 통해 협



<그림 5> 카네기멜론 대학도서관 데이터 저장 안내 페이지

*출처: <http://www.cmu.edu/research/data-management/storage-backup/index.html>. [cited 2014.9.14].

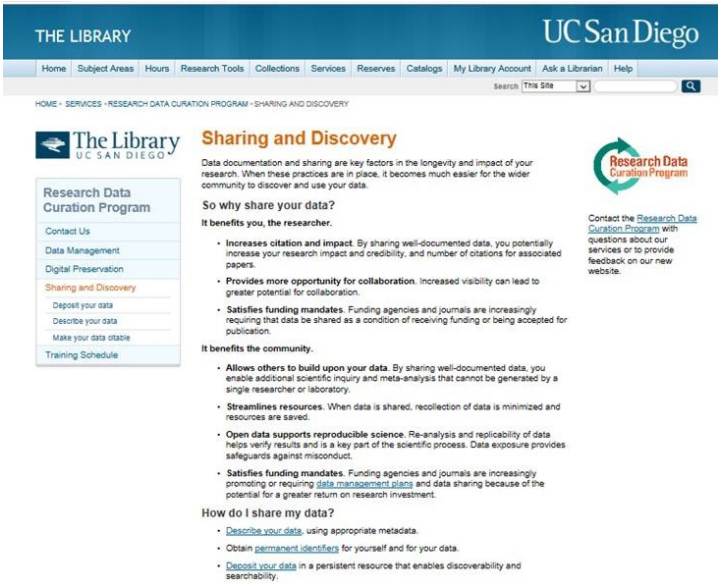
www.kci.go.kr

력연구나 공저자 참여의 기회 등을 기대하기도 한다(Sayogo and Pardo 2013). 조사결과 12개 기관(38.7%)에서 데이터 공유와 접근에 대한 설명을 제공하고 있었으며 데이터 공유의 필요성과 공유방법, 공유와 관련된 이슈 및 예외사항을 제시하고 있었다.

데이터 공유의 필요성은 5개 기관에서 제시하고 있었으며 그 중 듀크 대학도서관의 예를 살펴보면 (1) 연구비지원기관의 의무사항이므로 이를 준수; (2) 데이터의 공유를 통해 연구의 투명성과 재현성을 증가; (3) 데이터의 관리와 보존을 통해 연구의 연속성 유지; (4) 데이터의 인용을 통해 연구의 가시성을 증가할 수 있다는 점을 언급하였다

데이터의 공유방법은 5개 기관에서 언급되었는데 캘리포니아 샌디에고 대학도서관에서는 그 절차를 데이터 기술, 데이터에 대한 영구적인 ID 제공, 데이터 기탁의 세 단계로 설명하였다(〈그림 6〉 참조). 하버드대학도서관의 경우 데이터를 공유하기 전 체크리스트로서 (1) 데이터를 공유하기에 적합한 레포지터리 선택; (2) 데이터 재이용을 위해 데이터에 대한 기술을 명확히 할 것; (3) 타인에 의해 분석이 가능한 형태로 데이터를 공유할 것; (4) 민감한 정보에 대해서는 익명화를 실시할 것; (5) 논문을 출판하였을 때 사용했던 분석방법으로 같은 결과가 나오는지 확인해 볼 것; (6) 데이터에 개인정보나 민감한 정보가 있을 경우 공개를 지연하는 엠바고(embargo) 기간을 어느 정도로 할 것인지를 정할 것을 제시하였다.

데이터 공유를 둘러싼 이슈나 예외사항에 대해서는 4개 기관에서 이를 제시하였다. 펜실베



〈그림 6〉 캘리포니아 샌디에고 대학도서관 데이터 공유 안내 페이지

*출처: 〈http://isites.harvard.edu/icb/icb.do?keyword=k78759&tabgroupid=icb.tabgroup124911〉. [cited 2014.9.14].

www.kci.go.kr

이니아 주립 대학도서관의 경우 데이터 공유를 통해 발생할 수 있는 개인정보보호 침해 등 윤리적인 문제를 고려할 것과, 데이터의 접근수준을 공개, 엠바고기한 제공, 접근제한 중에서 어느 수준으로 결정할 것인지를 고려해야 한다고 제안하였다. 스탠포드대학도서관에서는 데이터가 다른 소스로부터 추출된 것일 경우 데이터의 통제권이 누구에게 있는지를 판단하기 어려운 경우가 있으며 공유 전에 이를 확인해야 함을 지적하였다. 또한 저작권, 특허, 상표권 등 지적재산권과 관련된 사항을 명확히 하고 크리에이티브 커먼스 라이선스를 활용한 데이터 라이선싱의 방안을 제시하였다.

3.6 데이터 보존

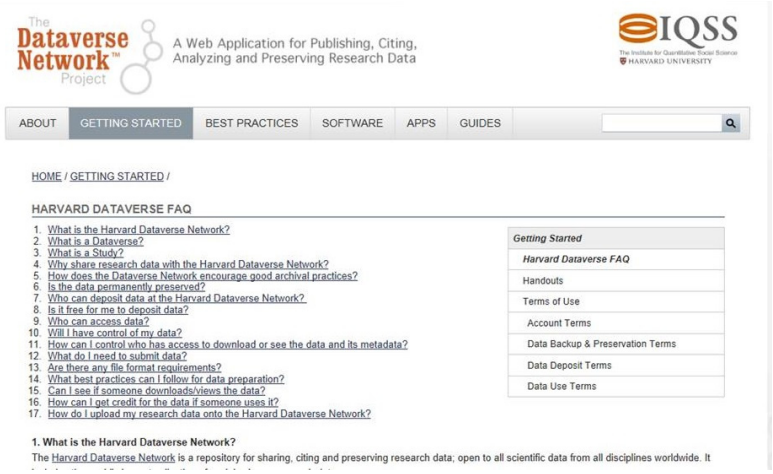
RDM 서비스의 중요한 기능 중의 하나는 가치있는 데이터를 선정, 평가하여 이를 보존하는 것이며 Moon(2014)은 캐나다 퀸즈 대학(Queen's University) 도서관의 사례를 통해 데이터 보존 기능을 RDM 서비스의 향후 비전으로 제시하였다. 조사대상 도서관에서는 구체적인 데이터 보존의 업무를 명시하기보다는 대부분 보존을 수행하는 레포지터리로의 안내를 제공하고 있었으며 18개 기관(58.1%)에서 교내 혹은 외부의 레포지터리를 제시하고 있었다.

이 중 13개 기관에서는 교내에서 제공되는 기관 레포지터리 또는 디지털도서관을 데이터 공유와 보존의 장소로 소개하고 있었다. 예를 들어 인디애나 블루밍턴 대학의 경우 'IU ScholarWorks'라는 기관 레포지터리를 전반적인 RDM 서비스 제공의 플랫폼으로 활용하고 있었으며 데이터의 보존도 이를 통해 수행하고 있었다. 럿거스 대학에서도 'RUcore'라는 기관 레포지터리를 데이터 보존의 장소로 제시하고 있었으며 터프츠 대학에서는 'Tufts Digital Library'라는 시스템을 통해 교내 연구자들의 데이터를 수집, 보존하고 공유하는 업무를 수행하고 있었다.

데이터 보존과 공유를 위한 외부 레포지터리는 분야별로 매우 다양한 양상을 보이고 있었으며 이를 반영하듯 10개 기관이 연구데이터 레포지터리 검색 시스템인 Databib를 제시하여 연구자들이 직접 원하는 데이터 레포지터리를 선택하도록 유도하고 있었다. 다양한 레포지터리 중에서도 11개 기관이 사회과학분야 데이터를 보존, 공유하는 ICPSR 데이터 아카이브로의 안내를 제공하고 있었다. 또한 9개 기관에서 생물학분야의 학술논문 및 데이터 레포지터리인 Dryad를 제시하고 있었다. 조지아 대학도서관의 경우 다양한 외부 레포지터리를 학문분야별로 농학, 고고학, 우주과학, 대기과학, 생명과학, 생태학, GIS, 지구과학, 해양학, 물리학, 사회과학, 기상학으로 구분하여 링크를 제공하고 있었다.

데이터 보존에 특화된 레포지터리 서비스로 하버드 대학에서 개발한 Harvard Dataverse Network이 있으며 조사대상기관 중 하버드대학도서관을 제외한 3개 기관에서도 이를 외부 레포지터리로서 안내하고 있었다. 이는 모든 학문분야의 연구데이터를 국제적인 범위에서 수집한다는 목표에 기반을 두고 운영되고 있다. 데이터 보존을 위해 보존메타데이터와 XML 메타데이터를 제공하며 여러 장소에 데이터를 복사하여 보관하는 LOCKSS(Lots of Copies Keep Stuff Safe)기법을 활용하고 있다(〈그림 7〉 참조).

www.kci.go.kr



〈그림 7〉 하버드 대학도서관 Dataverse Network 안내 페이지

*출처: 〈http://thedata.org/book/faq-using-harvard-dataverse-network〉. [cited 2014.9.14].

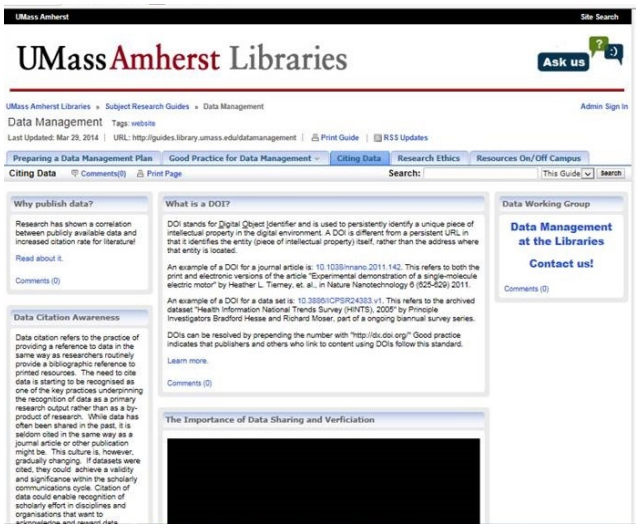
3.7 데이터 인용

데이터의 공유에 있어 연구자들이 우려하는 부분 중 하나는 타인에 의해 데이터가 제대로 인용되지 않은 채 표절될 수 있는 등 데이터가 부당하게 이용될 수 있다는 점이다(Campbell 외 2002). 효과적인 데이터의 재이용을 위해서는 적절한 데이터 인용이 필수적이므로 RDM 서비스에서 이에 대한 안내와 교육을 제공하는 것이 필요하다. 북미의 223개 대학도서관을 중심으로 RDM 서비스를 설문조사한 연구에서도 데이터 인용 안내가 가장 일반적으로 제공되는 서비스 요소인 것으로 나타났다(Tenopir 외 2014).

조사대상기관 중에서는 15개 기관(48.4%)에서 데이터 인용 안내를 제공하는 것으로 나타났다. 데이터 인용에 대한 내용은 크게 인용의 필요성, 인용의 요소와 예시, 그리고 DOI(Digital Object Identifier)제공에 관한 것이다.

데이터 인용의 필요성은 7개 기관에서 제시되었는데 매사추세츠 앰허스트 대학도서관의 경우 데이터가 더 이상 연구 활동의 부산물이 아니라 중요한 연구결과물로서 인식되고 있으므로 연구자들의 노력을 인정하고 보상한다는 차원에서 데이터에 대한 인용이 매우 중요하다고 밝히고 있다(〈그림 8〉 참조). 데이터 인용의 요소를 제시하는 기관은 모두 8개 기관이었는데 이 중 하버드 대학도서관에서 제시하는 인용의 요소를 살펴보면 저자명, 제목, 공개 혹은 배포날짜, 데이터가 공개된 레포지터리, 데이터의 판(edition)/버전/권(volume), 분석에 사용된 소프트웨어, URI나 고유 식별자, 접근날짜 등 접근정보가 있다. 특히 DOI와 같은 고유 식별자를 데이터에 제공함으로써 데이터의 인용을 용이하게 할 수 있으며, 5개 기관에서 데이터의 DOI를 등록할 수 있는 서비스인 DataCite에 대한 안내를 제공하고 있었다.

www.kci.go.kr



〈그림 8〉 매사추세츠 앰허스트 대학도서관 데이터 인용 안내 페이지

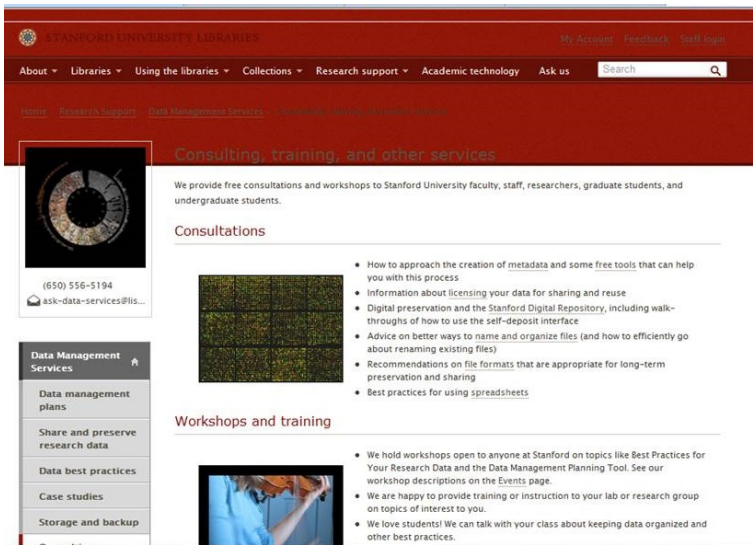
*출처: 〈http://guides.library.umass.edu/content.php?pid=258091&sid=2147156〉. [cited 2014.9.14].

3.8 데이터관리 교육

데이터의 공유 및 공개가 강조되면서 모든 연구자들이 데이터관리 리터러시(data management literacy), 즉 데이터를 관리, 검색, 보존하는 방법과 절차를 이해하고 활용하는 능력이 필요하며 이에 대한 도서관의 역할이 더욱 강조되고 있는 실정이다(Haendel 외 2012). 조사대상기관 중에서도 다수를 차지하는 20개 기관(64.5%)에서 데이터관리에 대한 개별 컨설팅 서비스, 워크숍, 튜토리얼 등 교육자료 제공 가운데서 하나 이상의 서비스를 제공하고 있었다. 이들 중에서 11개 기관은 개별 컨설팅 서비스와 워크숍을 함께 제공하고 있었으며 교육자료 제공까지 함께 실행하는 3개 기관도 이에 포함되어 있는 것으로 나타났다. 2개 기관은 워크숍만 제공하고 있었으며 다른 2개 기관은 교육 자료만을 제공하고 있었다.

교육 내용에 대한 정보를 제공하는 기관은 10개 기관이었으며 이들 중에서 인디애나 블루밍턴 대학도서관과 휴스턴 대학도서관은 컨설팅 또는 워크숍을 통해 DMP 작성에 대한 교육만을 제공하고 있다고 언급하였다. 그러나 나머지 8개 기관의 경우 DMP 작성과 관련된 내용과 함께 데이터 관리에 대한 전반적인 사항을 교육하고 있었다. 예를 들어 프린스턴 대학도서관의 경우 개별 컨설팅을 통해 DMP 작성을 지원하며 개인 또는 그룹 단위로 RDM에 대한 교육을 제공하고 기관 레포지터리 참여나 분야별 데이터 레포지터리 선택에 관해서도 담당자에게 연락할 경우 개별적인 컨설팅을 제공하고 있었다. 스탠포드 대학도서관에서도 개별 컨설팅을 통해 메타데이터, 데이터 파일 정리, 레포지터리 선택, 라이선싱 등 전반적인 RDM에 대한 교육을 제공하는 것으로 나타났다(〈그림 9〉 참조).

www.kci.go.kr



〈그림 9〉 스탠포드 대학도서관 데이터관리 교육 안내 페이지

*출처: 〈http://library.stanford.edu/research/data-management-services/consulting-training-and-other-services〉. [cited 2014.9.14].

3.9 데이터 지적재산권

조사대상기관 중 9개 기관(29%)에서 데이터의 지적재산권 및 라이선싱과 관련된 내용을 언급하고 있었다. 미네소타대학 도서관과 럿거스 대학 도서관에서 제공하는 저작권 관련 설명에 따르면 미국의 저작권법에서 연구데이터는 저작물이 아닌 사실(facts)로 간주되므로 저작권을 주장할 수 없다. 그러나 데이터베이스 내의 데이터 일부나 표 또는 그래프로 표현된 데이터일 경우 저작권을 주장할 수도 있다.

조사 결과 9개 기관 중 8개 기관에서 교내 혹은 외부의 관련 정책이나 안내를 제공하고 있었으며 4개 기관에서 크리에이티브 커먼스 라이선스 등 데이터에 적용 가능한 라이선스에 대한 정보를 제공하고 있었다. 특히 데이터를 모든 사람들이 어떤 목적으로도 이용할 수 있게 공적 영역에 있는 자료로서 공개하고자 할 때 CC0 (Creative Commons Zero) 라이선스를 사용할 수 있음을 제시하고 있다.

4. 결론 및 제언

본 연구에서는 아직 초기단계에 있지만 미국 대학도서관을 중심으로 점차 확산되어가는 RDM 서비스의 구성요소와 서비스의 내용을 도서관 웹사이트에 공개되어 있는 정보를 바탕으로 분석하였다. 조사대상이 된 31개 연구중심대학도서관 가운데 대다수의 기관에서 DMP 작성지원 서비스를 제공하고 있었으며 절반 이상의 기관에서 데이터 기술, 데이터 보존, 데이터 관리 교육을 실시하고 있었다. 또한 데이터 인용 안내도 절반에 가까운 조사대상기관에서 제공하

www.kci.go.kr

고 있었다. 이는 미국의 대학도서관에서 DMP 작성지원이라는 좁은 범위의 서비스를 넘어 데이터의 관리와 보존을 포괄하는 RDM 서비스를 구축하고 있음을 나타낸다. 데이터 기술과 데이터 보존에 있어서는 연구자들의 학문분야별로 적절한 메타데이터 표준을 제시, 안내하고 보존을 위한 레포지터리 선정을 안내하는 것에 초점을 맞추고 있었다. 또한 다수의 기관에서 워크숍 등 교육 프로그램뿐만 아니라 개별 컨설팅 서비스를 제공하면서 연구자들의 데이터 관리 요구에 적극적으로 대응하는 모습을 보이고 있었다.

국내의 KISTI P-CUBE나 연구재단의 기초학문자료센터 등 과학데이터 또는 인문사회분야 데이터를 수집, 공유, 보존 및 활용하려는 서비스에서는 본 연구에서 제시된 RDM 서비스의 구성요소와 내용을 서비스 개선을 위한 기초자료로서 활용할 수 있을 것이다. 또한 본 연구의 분석내용을 기반으로 향후 국내 RDM 서비스 개발을 위한 시사점을 제시하면 다음과 같다.

첫째, RDM 서비스는 기존 도서관 조직과 직원의 역량만으로는 운영되기 어려우며 기관 내 IT 서비스 기관, 연구처, 법/정책 관련 부서 등 다양한 기관과의 협력을 통해 효과적인 서비스가 제공될 수 있다. 미국 내에서도 RDM 서비스의 복잡성과 많은 자원 및 인력이 요청된다는 점에서 서비스의 지속가능성에 대한 우려가 존재한다. 따라서 국내의 RDM 서비스 운영을 위한 협력방안의 모색과 이를 효과적으로 지속할 수 있는 전략에 대한 논의가 필요할 것이다. 둘째, RDM 서비스를 효율적으로 운영하기 위한 사서의 지식과 역량 및 이들을 양성하기 위한 교육 프로그램에 대한 논의가 필요하다. 데이터 관리가 대학도서관 사서의 업무 영역에 포함되기 시작한 미국이나 영국의 경우 Data Librarian 등 RDM 서비스 담당사서의 채용이 증가하면서 이에 대한 논의가 더욱 활발히 진행되고 있다. 국내에서도 RDM 서비스의 구축이 진행되는 상황에서 서비스 담당자가 갖출 지식과 기술 및 자질 및 향후 이러한 역량을 갖춘 전문 인력을 양성하는 방안에 대하여 지속적인 논의와 연구가 필요할 것이다.

미국의 RDM 서비스는 학술커뮤니케이션과 디지털보존, 기관 레포지터리 서비스, 오픈액세스 등 다양한 주제들과 접점을 가지고 발전하고 있다. 국내에서도 도서관 역할의 변화를 주도하는 서비스로서 RDM 서비스 구축과 제공이라는 실무적 측면과 함께 RDM 서비스를 둘러싼 다양한 이해관계자 간 상호작용을 분석하는 이론적 측면에 대한 연구가 함께 이루어져야 할 것이다.

참 고 문 헌

Adamick, J., R. C. Reznik-Zellen, and M. Sheridan. 2012. "Data Management Training for Graduate Students at a Large Research University." *Journal of eScience Librarianship*, 1(3): 180-188.

www.kci.go.kr

Campbell, E. G., B. R. Clarridge, M. Gokhale, L. Birenbaum, S. Hilgartner, N. A. Holtzman, and D. Blumenthal. 2002. "Data withholding in academic genetics: evidence from a national survey." *JAMA*, 287(4): 473-480.

Carnegie Foundation. 2014. *The Carnegie Classification of Institutions of Higher Education.* [online]. [cited 2014.9.14]. 〈http://classifications.carnegiefoundation.org/〉.

Cox, A. M. and S. Pinfield. 2013. "Research Data Management and Libraries: Current Activities and Future Priorities." *Journal of Librarianship and Information Science*, June 28, 2013. [online]. [cited 2014.9.14]. 〈http://dx.doi.org/10.1177/0961000613492542〉.

Fearon, D. Jr., B. Gunia, B. E. Pralle, S. Lake, and A. L. Sallans. 2013. *SPEC Kit 334: Research Data Management Services*. Washington, DC: ARL.

Ferguson, J. 2012. "Lurking in the Lab: Analysis of Data from Molecular Biology Laboratory Instruments." *Journal of eScience Librarianship*, 1(3). [online]. [cited 2014.9.14]. 〈http://dx.doi.org/10.7191/jeslib.2012.1019〉.

Haendel, M. A., N. A. Vasilevsky, and J. A. Wirz. 2012. "Dealing with Data: A Case Study on Information and Data Management Literacy." *PLoS Biology*, 10(5), e1001339. [online]. [cited 2014.9.14]. 〈http://www.plosbiology.org/article/info%3Adoi%2F10.1371%2Fjournal.pbio.1001339〉.

Hswe, P. and A. Holt. 2011. *NSF Data Sharing Policy*. [online]. [cited 2014.9.14]. 〈http://www.arl.org/focus-areas/e-research/data-access-management-and-sharing/nsf-data-sharing-policy〉.

Johnston, Lisa, Meghan Lafferty, and Beth Petsan. 2012. "Training Researchers on Data Management: A Scalable, Cross-Disciplinary Approach." *Journal of eScience Librarianship*, 1(2). [online]. [cited 2014.9.14]. 〈http://dx.doi.org/10.7191/jeslib.2012.1012〉.

Jones, S., A. Ball, and Ç. Ekmekcioglu. 2008. "The Data Audit Framework: A First Step in the Data Management Challenge." *International Journal of Digital Curation*, 3(2): 112-120.

Jones, S., G. Pryor, and A. Whyte. 2013. *How to Develop Research Data Management Services —A Guide for HEIs*. Edinburgh: Digital Curation Centre.

Lyon, L. 2012. "The informatics Transform: Re-engineering Lbraries for the Data Decade." *International Journal of Digital Curation*, 7(1): 126-138.

Molloy, Laura and Kellie Snow. 2012. "The Data Management Skills Support Initiative: Synthesising Postgraduate Training in Research Data Management." *International Journal of Digital Curation*, 7(2): 101-109.

Moon, J. 2014. "Developing a Research Data Management Service: A Case Study." *Partnership:*

www.kci.go.kr

the Canadian Journal of Library and Information Practice and Research, 9(1): 1-14.

Raboin, Regina, Rebecca C. Reznik-Zellen, and Dorothea Salo. 2012. "Forging New Service Paths: Institutional Approaches to Providing Research Data Management Services." *Journal of eScience Librarianship*, 1(3): 134-147.

Sallans, A. and M. Donnelly. 2012. "DMP Online and DMPTool: Different Strategies towards a Shared Goal." *International Journal of Digital Curation*, 7(2): 123-129.

Sayogo, D. S. and T. A. Pardo. 2013. "Exploring the Determinants of Scientific Data Sharing: Understanding the Motivation to Publish Research Data." *Government Information Quarterly*, 30: S19-S31.

Starr, Joan, Perry Willett, Lis Federer, Horning Claudia, and Mary Linn Bergstrom. 2012. "A Collaborative Framework for Data Management Services: The Experience of the University of California." *Journal of eScience Librarianship*, 1(2). [online]. [cited 2014.9.14]. 〈http://dx.doi.org/10.7191/jeslib.2012.1014〉.

Steeleworthy, M. 2014. "Research Data Management and the Canadian Academic Library: An Organizational Consideration of Data Management and Data Stewardship." *Partnership: the Canadian Journal of Library and Information Practice and Research*, 9(1). [online]. [cited 2014.9.14]. 〈https://journal.lib.uoguelph.ca/index.php/perj/article/view/2990〉.

Tenopir, C., R. J. Sandusky, S. Allard, and B. Birch. 2014. "Research Data Management Services in Academic Research Libraries and Perceptions of Librarians." *Library & Information Science Research*, 36(2): 84-90.

Verbaan, E., and A. M. Cox. 2014. "Occupational Sub-cultures, Jurisdictional Struggle and Third Space: Theorising Professional Service Responses to Research Data Management." *The Journal of Academic Librarianship*, 40: 211-219.

Witt, M. 2012. "Co-designing, Co-developing, and Co-implementing an Institutional Data Repository Service." *Journal of Library Administration*, 52(2): 172-188.

www.kci.go.kr

[부록] 31개 조사대상기관 RDM 서비스 웹사이트

기관명	URL
카네기멜론(Carnegie Mellon Univ.)	http://www.cmu.edu/research/data-management/
콜로라도주립(Colorado State Univ.)	http://lib.colostate.edu/services/data-management
코넬(Cornell Univ.)	http://data.research.cornell.edu/
듀크(Duke Univ.)	http://library.duke.edu/data/data-consulting
플로리다주립(Florida State Univ.)	http://www.lib.fsu.edu/tads/datamanagement
하버드(Harvard Univ.)	http://isites.harvard.edu/icb/icb.do?keyword=k78759
인디애나 블루밍턴 (Indiana Univ._Bloomington)	http://scholarworks.iu.edu/data/index.php
메사추세츠공과대(MIT)	http://libraries.mit.edu/data-management/
뉴욕(New York Univ.)	http://guides.nyu.edu/friendly.php?s=dataservices
펜실베이니아주립(Penn State Univ.)	http://www.libraries.psu.edu/psul/researchguides/pubcur/datatoolkit.html
프린스턴(Princeton Univ.)	http://library.princeton.edu/research-data-management
퍼듀(Purdue Univ.)	https://www.lib.purdue.edu/research
럿거스(Rutgers Univ.)	http://libguides.rutgers.edu/data
스탠포드(Stanford Univ.)	http://library.stanford.edu/research/data-management-services
테네시(Univ. of Tennessee)	http://libguides.utk.edu/datamanagement?hs=a
터프츠(Tufts Univ.)	http://researchguides.library.tufts.edu/datamanagement
뉴욕주립 버펄로(Univ. at Buffalo)	http://library.buffalo.edu/scholarly/data/
캘리포니아 로스엔젤레스(UCLA)	http://www.library.ucla.edu/service/data-management-curation-services
캘리포니아 샌디에고(UC San Diego)	http://libraries.ucsd.edu/services/data-curation/
신시내티(Univ. of Cincinnati)	http://guides.libraries.uc.edu/datamanagementplanning
조지아(Univ. of Georgia)	http://guides.libs.uga.edu/content.php?pid=489484&sid=4014496
하와이(Univ. of Hawaii)	http://guides.library.manoa.hawaii.edu/c.php?g=105213&p=685783
휴스턴(Univ. of Houston)	http://info.lib.uh.edu/services/faculty-and-graduate-students/data-management-resources
매사추세츠 앰허스트 (Univ. of Massachusetts_Amherst)	http://www.library.umass.edu/services/services-for-faculty/data-management/
마이애미(Univ. of Miami)	http://libguides.miami.edu/content.php?pid=174335
미네소타(Univ. of Minnesota)	https://www.lib.umn.edu/datamanagement
네브래스카(Univ. of Nebraska)	http://libraries.unl.edu/data-management
유타(Univ. of Utah)	http://www.lib.utah.edu/services/digital-scholarship-lab/data-management.php
워싱턴(Univ. of Washington)	http://guides.lib.washington.edu/data
위스콘신 매디슨 (Univ. of Wisconsin_Madison)	http://researchdata.wisc.edu/
버지니아 커먼웰스 (Virginia Commonwealth Univ.)	http://www.library.vcu.edu/services/data/

www.kci.go.kr

www.kci.go.kr