

# 외국 대학도서관의 디지털 큐레이션 프로세스 비교분석\*

## Comparative Analysis on Digital Curation Process in Foreign Academic Libraries

한 나 은(Na-eun Han)\*\*  
김 성 희(Seonghee Kim)\*\*\*

### < 목 차 >

I. 서론	IV. 대학도서관의 디지털 큐레이션 프로세스 분석
II. 이론적 배경	1. 데이터 선정 및 입수
1. 연구데이터	2. 보존 활동
2. 디지털 큐레이션	3. 접근 및 이용
3. 선행연구	4. 책임 및 관리
III. 디지털 큐레이션 모델 분석	5. 종합분석
1. DCC 생애주기 모델	V. 결론 및 제언
2. DCU 디지털 큐레이션 프로세스 모델	
3. DCC & U 생애주기 모델	

### 초 록

본 연구에서는 연구데이터의 효율적인 활용을 위해 국외의 10개 대학도서관의 보존정책을 대상으로 디지털 큐레이션 프로세스를 분석하였다. 분석은 데이터 선정 및 접수, 보존 활동, 접근 및 이용, 책임 및 관리의 4가지 부분으로 나누어 실시하였다. 분석결과 데이터 선정 및 접수에서 수집 대상 데이터는 대부분 대학 내에서 생산된 디지털 형태의 연구데이터나 대학에서 생산하지는 않았으나 대학을 통해 이용할 수 있는 디지털 데이터로 나타났다. 보존을 위해 사용되고 있는 메타데이터는 Dublin Core, MODS, PREMIS 등으로 나타났으며, 접근 및 이용에서는 법률 및 규약에 기초하여 접근, 규제 등을 결정하는 것으로 나타났다. 책임 및 관리 차원에서는 관리적, 재정적, 기술적, 물리적, 인적 부분 등으로 나누어 각각의 감사나 지속가능성 등을 검토하는 것으로 나타났다.

키워드: 디지털 큐레이션, 큐레이션 프로세스, 연구데이터, 대학도서관

### ABSTRACT

This study analyzed the digital curation process targeting the foreign 10 university libraries for the efficient utilization of the research data. The analysis was conducted by dividing into 4 parts such as the selection and ingest of data, preservation, access and use, responsibilities and management. The results indicated that research data for selection and ingest include data produced by the faculty members or can be accessible in the university or research institution. The findings also showed that the institutions use Dublin Core, MODS, and PREMIS as a metadata for preservation. In the aspect of approach and use, regulation are decided based on the law and agreement. And in the aspect of responsibility and management, most elements are divided into the management, financial, technical, physical, and personal parts and each audit or sustainability is reviewed.

Keywords: Digital curation, Curation process, Research data, Academic libraries

\* 이 논문은 2014년도 중앙대학교 연구장학기금 지원에 의한 것임.

\*\* 중앙대학교 문헌정보학과 박사과정(euni\_zz@naver.com)

\*\*\* 중앙대학교 문헌정보학과 교수(seonghee@cau.ac.kr)

• 논문접수: 2014년 5월 21 • 최초심사: 2014년 5월 22일 • 게재확정: 2014년 6월 11

## I. 서론

인터넷 환경의 변화 및 정보통신기술의 발달로 디지털 자료의 양이 급속하게 팽창하였고, 디지털 정보자원에 대한 접근법도 PC에서 스마트폰까지 점차 다양해지고 있다. EMC(Enterprise Content Management)가 IDC(International Data Corporation)에 의뢰하여 2011년 12월에 발표한 'IDC Digital Universe Study'에 따르면 전세계 디지털 데이터 양은 2020년이면 40제타 바이트(ZB)에 도달할 것이라고 전망하고 있다.<sup>1)</sup> 디지털 데이터의 양적 증가에 따른 체계적 관리, 그와 더불어 훼손되거나 소실되기 쉬운 디지털 자료의 특성상 장기적 보존의 필요성이 증대되었다. 디지털 자료들은 연속성을 보장하기 위해 생애주기에 맞추어 관리해야 할 필요가 있다. 디지털 자료들은 생성, 관리, 보존, 활용, 그리고 폐기되는 일련의 과정을 거치게 되는데, 이 과정들에 대한 정확한 이해가 수반되어야 효과적인 큐레이션 및 생애주기 프로세스에 따른 각 단계별 큐레이션 서비스를 실시할 수 있다.

특히, 대학 및 연구소에서 생산되는 연구결과 및 연구 수행과정에서 생산되는 데이터는 연구수행과정에 대한 기록으로 연구결과에 대한 핵심정보를 담고 있기 때문에 연구결과를 증명할 수 있는 중요한 가치를 지니고 있다. North Carolina State University,<sup>2)</sup> Illinois State University,<sup>3)</sup> University of Michigan<sup>4)</sup> 등 국외의 대학 및 도서관에서 이미 디지털 큐레이션에 관해 연구를 진행하고 있으며 대학 내 교과과정에서도 디지털 큐레이션과 관련한 프로그램이 개설되고 있다. 그러나 국내에서는 디지털 큐레이션과 관련한 연구 및 논의가 미흡한 편이다.

국내에서는 최근들어 연구데이터 또는 디지털 큐레이션에 관한 연구가 진행되었는데 우선, 연구데이터와 관련해서는 연구데이터 수집방안, 생명주기를 기반으로 한 정책 제안이나 디지털 큐레이션 모형설계에 관한 연구가 진행되어 왔다<sup>5)</sup>. 디지털 큐레이션 사례분석과 관련한 연구로는 영국의 DCC와 같은 대표적인 기관 한 개를 대상으로 연구가 진행되어 왔다.<sup>6)</sup> 하지만 아직까지 대학기관을 대상으로 전반적인 디지털 큐레이션 프로세스를 구체적으로 비교한 연구는 없는 실정

1) EMC & IDC, IDC Digital Universe Study: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East, <<http://www.emc.com/collateral/analyst-reports/idc-digital-universe-united-states.pdf>> [cited 2014. 5. 7].

2) University of North Carolina at Chapel Hill, Certificate in digital Curation, <[http://sils.unc.edu/programs/certificates/digital\\_curation](http://sils.unc.edu/programs/certificates/digital_curation)> [cited 2014. 6. 8].

3) University of Illinois at Urbana-Champaign, Data Curation Education Program, <<http://www.lis.illinois.edu/research/projects/data-curation-education-program>> [cited 2014. 6. 8].

4) University of Michigan, Master of Science in Information, Preservation of Information specialization, <<http://www.si.umich.edu/academics/msi/preservation-information-pi>> [cited 2014. 6. 8].

5) 김은정, 남태우, "연구데이터 수집에 영향을 미치는 요인 분석," 정보관리학회지, 제29권, 제2호(2012), p.28; 안영희, 박옥화, "디지털 큐레이션 정책을 위한 프레임워크 개발," 한국도서관·정보학회지, 제41권, 제1호(2010), pp.167-186.

6) 설문원 등, 국가 디지털 아카이빙 체제 구축에 관한 연구(서울 : 한국과학기술정보연구원, 2005).

이다. 따라서 본 연구에서는 연구데이터의 효율적인 보존 및 활용을 위해 대학도서관을 중심으로 디지털 큐레이션 프로세스 및 보존정책을 분석하였다. 이를 위해 디지털 보존정책을 성문화하여 수립 및 시행하고 있는 국외 10개의 대학도서관을 선정하여 각각의 디지털 큐레이션 프로세스 및 보존정책 분석을 실시하였다. 분석 기준 설정을 위해 기관별 혹은 연구자별로 상이하게 정의되어 있는 디지털 큐레이션 프로세스를 분석하여 공통 요소를 추출하였으며, 앞서 선정한 국외 대학도서관의 디지털 큐레이션 프로세스 및 보존정책 분석 결과를 바탕으로 향후 국내 대학도서관의 디지털 큐레이션 서비스 도입 방안에 대해 제안하였다.

## II. 이론적 배경

### 1. 연구데이터

연구자들의 연구 행위 또는 활동 과정에서 연구의 부산물이자 결과물로서 각종 데이터가 생산되고 있는데 현재 연구데이터라는 용어에 대해서는 문헌정보학 용어사전이나 기록관리학 용어사전에 정의가 되어있지 않다. 명문화된 지침으로 연구데이터에 대해서 규정한 사례를 살펴보면 서울대학교 연구윤리지침과 OECD(Organization for Economic Co-operation and Development)의 정의가 있다.

「서울대학교 연구윤리지침」에서는 연구데이터를 “실험의 재료나 과정 및 결과, 관찰이나 현장 조사 및 설문조사의 결과 등의 원자료(raw data)를 의미한다. 연구자료는 연구데이터 및 이를 처리한 2차 자료로 다른 연구자가 동일한 조건하에 동일한 결과를 재현할 수 있도록 정해진 방법에 따라서 명확히 기록해야 한다”고 정의하였다.<sup>7)</sup>

OECD에서 연구데이터는 “수치점수, 문자기록, 이미지나 소리 등의 사실적 데이터로서 주로 과학연구의 주요한 원천으로 사용하며, 과학단체에서 연구결과를 검증하기 위해 필수적인 것으로 규정한다”라고 정의되어 있다.<sup>8)</sup>

국내·외 연구자들이 연구데이터에 대한 정의를 시도하였는데, 먼저 김선태 등은 과학데이터라는 용어를 사용하였으며, 과학데이터를 “연구자의 연구 활동 과정 중 생성되는 다양한 유형의 사실적 기록을 의미한다. 즉, 연구활동을 통하여 생산된 연구활동의 기록물로서 관측, 감시, 조사,

7) 서울대학교, 서울대학교 연구윤리지침, <[http://nrf.re.kr/nrf\\_tot/cms/include/download\\_general.jsp?file=d\\_0204\\_63513\\_0\\_98470.pdf](http://nrf.re.kr/nrf_tot/cms/include/download_general.jsp?file=d_0204_63513_0_98470.pdf)> [인용 2014. 5. 7].

8) OECD, OECD principles and guidelines for access to research data from public funding, <<http://www.oecd.org/science/sci-tech/38500813.pdf>> [cited 2014. 5. 7].

#### 4 한국도서관·정보학회지(제45권 제2호)

실험, 분석, 계산 등의 과정을 통하여 생산된 문자, 이미지, 오디오, 동영상 등의 아날로그 및 디지털 형식을 포괄하는 데이터”라고 정의하였다.<sup>9)</sup>

미국 National Academy of Sciences에서는 연구데이터를 “과학적, 공학적, 의학적 연구 분야에서 사용되는 정보이며, 연구의 결론을 이끌어내는 투입, 결과”로 정의하였다.<sup>10)</sup>

김은정은 연구데이터를 “연구수행과정 중에 산출되는 데이터로서 원데이터와 분석된 2차 자료를 포함하며, 관찰, 조사, 실험, 경험에 기반하는 사실이다. 이러한 데이터는 수치, 텍스트, 오디오, 컴퓨터 파일 등 다양한 형태의 데이터를 포함한다”라고 정의하였다.<sup>11)</sup>

이와 같이 연구데이터는 OECD의 지침과 같은 과학연구의 관점에서 보는 경우도 있으나, 공통적으로 학문분야와 관계없이 연구의 기초가 되는 다양한 유형의 데이터로 보고 있다. 그리고 이러한 데이터에 대해서는 연구과정에서 다양한 행위에 의해 산출되는 원자료로 규정하고 있음을 알 수 있으며 연구성과물인 연구보고서나 논문과는 구분된다.

따라서 본 연구에서는 이상의 여러 학자와 기관의 정의들을 통합하여 연구데이터란 “연구를 위해 사용된 실험 재료, 과정, 결과, 관찰, 설문조사의 원자료와 이를 이용해 처리된 2차 자료를 의미하며, 데이터의 형태는 문자, 수치, 이미지, 오디오, 동영상 등을 포함한다”고 정의하여 사용 하였다.

## 2. 디지털 큐레이션

디지털 큐레이션(Digital Curation)은 디지털 자원을 수집, 보존, 아카이빙, 제공하는 것을 지칭한다. 넓게 보면 현재와 장래에 이용될 신뢰할 수 있는 디지털 정보를 유지하고, 가치를 부여하는 것이라고 해석할 수 있다.<sup>12)</sup> 디지털 큐레이션은 연구자, 과학자, 역사학자 등이 현재와 장래에 참고할 수 있도록 디지털 자원의 장기 보존소를 설립하고 개발하여 재활용 하는 모든 활동을 포함한다는 점에서 디지털 아카이빙보다 넓은 범주의 활동을 지칭한다고 볼 수 있다.<sup>13)</sup>

디지털 큐레이션은 장·단기적인 관점에서 다양한 이점을 갖고 있다. 먼저 단기적으로는 신뢰할 수 있는 디지털 데이터에 대해서 지속적인 접근을 가능하게 해주며, 데이터 자체의 품질을 향상시켜주고 연구 성과물의 우수성을 높여 줄 수 있다. 또한 상호 검색 및 협력 증진을 위해 서로 다른 데이터 집합에서 공통 표준을 사용할 수 있도록 하며, 진위 테스트를 통해 데이터의 신뢰성을 향

9) 김선태 등, “과학데이터 보존 및 활용모델에 관한 연구,” 한국비블리아학회지, 제21권, 제4호(2010), p.84.

10) National Academy of Sciences, *Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age, Ensuring the integrity, accessibility, and stewardship of research data in the digital age*, (Washington, D.C.: National Academy Press, 2009), p.22.

11) 김은정, 연구데이터 수집에 영향을 미치는 요인 분석(박사학위논문, 중앙대학교 대학원 기록관리학과, 2012), p.16.

12) Digital Curation Centre, <<http://www.dcc.ac.uk/about/what/>> [cited 2014. 5. 7].

13) 국립중앙도서관 도서관연구소(KRILI), “도서관용어해설,” 도서관연구소 웹진, 제32권(2009), p.1.

상시킬 수 있게 해준다. 디지털 큐레이션을 통해 데이터의 진본성을 보장함으로써 법적으로 유효한 공식 기록으로 보존이 가능하고, 데이터를 사용한 이 후 재활용(re-use) 할 수 있게 됨으로써 정보의 재정적 가치를 보호하고 연구의 초기 투자에 있어서도 가치를 갖게 되는 것이다.

뿐만 아니라 장기적으로는 디지털 큐레이션을 통해 데이터를 보존하고, 손실이나 퇴화와 같은 위험으로부터 디지털 자료를 보호할 수 있고, 단기적인 재정적 변화나 제도적 변화에도 불구하고 데이터에 대한 지속적인 접근을 가능하게 해준다. 또한 데이터가 제공되는 맥락과 더불어 출처에 대한 정보가 제공되며, 데이터, 메타데이터 및 기타 정보 등 새로운 매체로 변환하는 도구 및 서비스를 제공함으로써 이용자들에게 유용한 서비스를 제공할 수 있다.<sup>14)</sup>

### 3. 선행연구

국내에서는 디지털 큐레이션이라는 용어를 직접 사용하여 진행된 연구들은 최근에서야 진행되고 있으며, 디지털 보존, 디지털 장기보존, 디지털 아카이빙 등의 용어와 혼재되어 사용되는 경우가 많다. 연구데이터와 관련해서도 연구데이터, 과학데이터, 연구기록물, 디지털 데이터 등 다양한 표현을 사용한 연구들이 진행되어왔다.

안영희, 박옥화는 영국 디지털 큐레이션 센터 및 8개 연구재단의 디지털 큐레이션 정책과 지원 서비스 사례를 통하여 디지털 정보자원의 생명주기 단계에 따른 큐레이션 정책 프레임워크를 개발하였다.<sup>15)</sup> 안영희, 박옥화는 또한 생명주기를 기반으로 한 생산자, 큐레이터, 이용자의 디지털 큐레이션 관계 모형을 근간으로 하여, 기관 레포지터리인 대학도서관의 디지털 자원에 대한 보존 및 큐레이션 모형을 설계한 바 있다.<sup>16)</sup>

설문원 등은 한국과학기술정보원의 데이터베이스 아카이빙 활동을 영국의 DCC(Digital Curation Centre) 활동 사례를 바탕으로 연구하였으며, 디지털 데이터에 대한 장기적 접근과 보존을 위한 전략계획 수립의 필요성을 주장하였다.<sup>17)</sup>

김선태 등은 기록물로서 과학데이터의 보존 및 활용 모델을 제시하였는데, 이를 위해 과학데이터와 관련하여 미국과 영국, 호주, 유럽연합의 국가별 동향을 조사하였다.<sup>18)</sup>

현문수는 디지털 연구데이터의 보존을 통해 얻을 수 있는 편익을 확인하기 위해서 경제학적 측면에서 디지털 연구데이터의 보존을 바라보고 있는 연구 사례를 살펴봄으로써 양적인 측면과

14) Abbott, D., "What is Digital Curation?," *DCC Briefing Papers: Introduction to Curation*, Edinburgh: Digital Curation Centre(2008), p.1.

15) 안영희, 박옥화, 전계논문, pp.167-186.

16) 안영희, 박옥화, "대학도서관 서비스의 디지털 큐레이션 전략," 한국도서관·정보학회지, 제40권, 제4호(2009), pp.311-326.

17) 설문원 등, 전계서.

18) 김선태 등, 전계논문, pp.81-93.

질적인 측면에서 디지털 연구데이터 보존을 통해 얻을 수 있는 광범위한 혜택을 분석하였다.<sup>19)</sup>

국외에서는 국내에 비해 디지털 큐레이션에 관한 연구가 활발히 진행되고 있으며, 디지털 보존의 중요성 증대와 함께 앞으로의 적용 방안에 관한 논의가 진행되어왔다.

먼저 Heidorn은 도서관이 갖는 자료에 대한 수집, 관리, 보존 및 폐기의 기능에 대해 접근하였으며, 더 이상 이러한 기능이 물리적 자료를 중심으로 강조될 것이 아니라 디지털 자료에도 중점을 두어야 하며 이를 위해 도서관의 역할이 데이터 큐레이션을 기반으로 하여 변화해야한다고 주장하였다.<sup>20)</sup>

Tyler는 The Georgia Institute of Technology(GT)의 발전과정을 연구하면서 미국 대학들에 적용할 수 있는 데이터 큐레이션 프로그램 개발을 위한 모델을 제안하였고,<sup>21)</sup> Jones는 영국 주요 연구기관들의 홈페이지 분석을 통해 각 기관의 정책과 데이터 관리에 대해 파악하여 데이터 관리와 공유의 필요성을 바탕으로 디지털 큐레이션을 연구의 초기 단계부터 도입해야 함을 주장하였다.<sup>22)</sup>

Jubb은 디지털 큐레이션의 수집대상인 저널, 그래프, 회의자료, 논문, 데이터, 회색문헌 등과 같은 다양한 형태의 연구데이터의 효율적인 활용을 위해 일관된 디지털 큐레이션 정책 프레임워크 개발이 필요함을 강조하였고,<sup>23)</sup> Uribe & Macdonald는 대학도서관이 21세기에 도래하며 직면하게 된 새로운 변화에 맞추어 적응하기 위해서는 사서가 데이터 큐레이터의 기능을 갖추어야 한다고 주장하였다.<sup>24)</sup>

Ray는 영국의 DCC와 미국의 IMLS(Institute of Museum and Library Services)의 디지털 큐레이션 활동에 대해 논하며 리포지터리의 설립 필요성을 강조하였고,<sup>25)</sup> Yakel은 큐레이션의 형태를 미국의 NSF(National Science Foundation)와 ACLS (American Council of Learned Societies), 영국의 UKOLN(The United Kingdom Office for Library and Information Networking)을 중심으로 분석하였다.<sup>26)</sup>

19) 현문수, "디지털 연구데이터 장기보존의 편익에 대한 연구," 한국기록관리학회지, 제11권, 제1호(2011), pp.161-181.

20) Heidorn, P. Bryan, "The Emerging Role of Libraries in Data Curation and E-science," *Journal of Library Administration*, Vol.51, No.7(2011), pp.662-672.

21) Tyler O. Walters, "Data Curation Program Development in U.S. Universities: The Georgia Institute of Technology Example," *The International Journal of Digital Curation*, Vol.3, No.4(2009), pp.83-92.

22) Jones, Sarah, A Report on the Range of Policies for and Related to Digital Curation, DCC Policies report, Version.1.2, pp.1-34, 2009. [cited 2014. 5. 7].

23) Jubb, Michael, "UK Research Funders' Policies for the Management of Information Outouts," *International Journal of Digital Curation*, Vol.2, No.1(2007), pp.29-48.

24) Uribe, Lusi Martines, and Macdonald, Stuart, "A new role for academic librarians: data curation," *Profesinoal de la Informacion*, Vol.17, No.3(2008), pp.273-280.

25) Ray, Joyce, *Managing the Digital World: the Role of Digital Curation*(UK:Edinburgh, 2008).

26) Yakel, Elizabeth, "Archives and Manuscripts Digital Curation, OCLC Systems & Services," *International Digital Library Perspectives*, Vol.23, No.4(2007), pp.335-340.

이상에서 국·내외에서 수행된 선행연구를 분석한 결과 주로 디지털 큐레이션 프레임 워크, 디지털 보존의 필요성 및 디지털 큐레이션 프로그램등에 대한 연구들이 수행되어 왔다. 하지만 대학도서관을 대상으로 디지털 큐레이션 프로세스 및 보존정책을 분석한 연구는 거의 없는 실정이다. 따라서, 본 연구에서는 외국의 대학도서관 가운데 10곳을 선정하여 디지털 큐레이션 프로세스와 보존정책을 분석하였다. 특히 본 연구는 보존정책의 유·무 뿐만 아니라 보존정책을 기반으로 디지털 큐레이션 프로세스를 분석했다는 점에서 기존 연구와의 차별성이 있다고 볼 수 있다.

### Ⅲ. 디지털 큐레이션 모델 분석

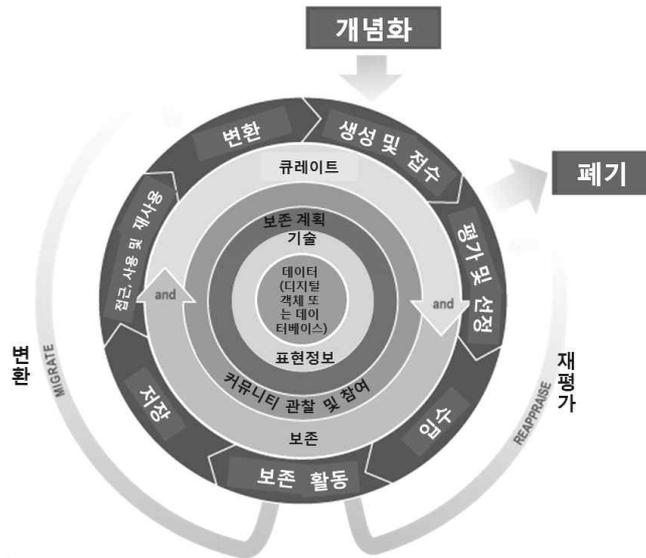
디지털 자료들은 생성, 관리, 보존, 활용, 그리고 폐기되는 일련의 과정에 대한 정확한 이해가 수반되어야 효과적인 큐레이션 및 생애주기 프로세스에 따른 각 단계별 큐레이션 서비스를 실시할 수 있다. 이 장에서는 고등 교육 연구 기관을 통해 연구 데이터 관리 및 보존 기술(technology) 구축 등에 초점을 맞추어 디지털 큐레이션에 대해 전문지식을 구축하고 있는 영국의 DDC에서 제공하는 생애주기 모델인 'DCC 생애주기 모델'과 Constantououlous & Dallas가 2008년 제안한 디지털 큐레이션 모델인 'DCU(Digital Curation Unit) 디지털 큐레이션 프로세스 모델', 그리고 DCC 생애주기 모델을 개선하기 위해 2009년 제안된 'DCC & U(Digital Curation Centre & Unit) 생애주기 모델'을 분석하여 공통요소를 추출하였다.

#### 1. DCC 생애주기 모델

DCC 생애주기 모델은 자료별로 세밀한 수준에 맞게 큐레이션 및 보존활동을 적절히 수행하기 위해서 개발된 것으로써, 큐레이션을 실시하는데 있어서 도움을 제공할 수 있는 관련 표준이라고 할 수 있다.<sup>27)</sup>

이 모델의 프로세스는 <그림 1>에서 보여주는 바와 같이 데이터, 개념화(Conceptualize), 생성 및 접수(Create and Receive), 평가 및 선정(Appraise and Select), 입수(Ingest), 보존 활동(Preservation Action), 저장(Store), 접근, 사용 및 재활용(Access, Use and Re-use), 변환(Transform)의 순서로 진행된다.

27) Sarah Higgins, "The DCC Curation Lifecycle Model," *The International Journal of Digital Curation*, Vol.1, No.3(2008), pp.135-136.



<그림 1> DCC 큐레이션 생애주기 모델<sup>28)</sup>

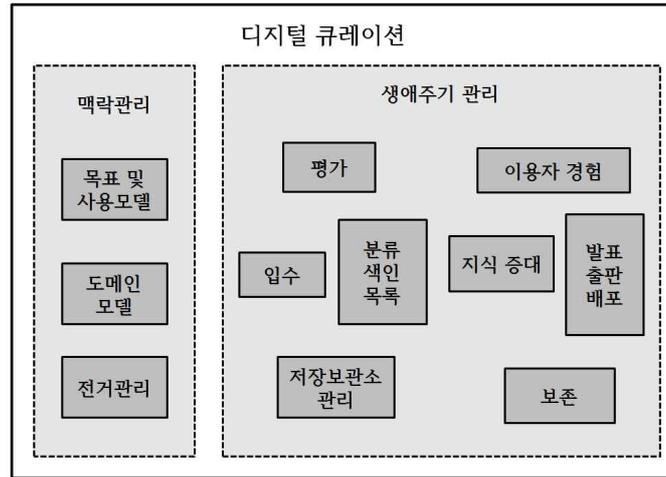
<그림 1>에서 보여주듯이 DCC 큐레이션 생애주기 모델은 초기 개념화에서 데이터를 성공적으로 큐레이션 및 보존하는 데 있어서 필요한 단계를 시각화를 통해 제공하고 있다. 이 모델은 큐레이션이 알맞은 순서로 필요한 모든 단계를 통해 수행될 수 있도록 보장하기 위해 조직이나 컨소시엄 내에서 활동을 계획하는데 사용될 수 있다. 뿐만 아니라 이 모델은 디지털 자료를 조직하고 관리하기 위해서 데이터 생산자, 데이터 큐레이터, 데이터 이용자들을 교육하는 도구로 이 모델을 사용하며, 뿐만 아니라 DCC의 디지털 자산 관리 정책 및 위험으로부터 보호 전략을 구축하는 데 있어서도 이 모델이 이용되고 있다.<sup>29)</sup>

## 2. DCU 디지털 큐레이션 프로세스 모델

Constantinou & Dallas이 제안한 DCU는 디지털 자료의 조직, 관리, 장기 보존 및 부가가치 서비스 지원에 대한 신뢰성을 높이고, 디지털 자료의 활용을 위한 새로운 접근법을 제공하고자 디지털 생애주기 관리(Digital Resources Lifecycle Management)와 맥락 관리(Context Management)로 구분하여 제시하였으며 <그림 2>와 같다.

28) *Ibid.*, p.136.

29) *Ibid.*, pp.135-136.



〈그림 2〉 DCU 디지털 큐레이션 프로세스 모델<sup>30)</sup>

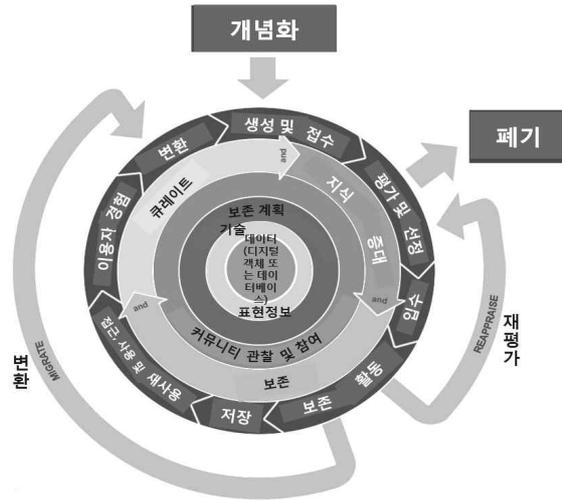
〈그림 2〉에서 나타난 것과 같이 디지털 생애주기 관리는 평가(Appraisal), 입수(Ingest), 분류, 색인 및 목록(Classification, Indexing and Cataloguing), 지식 증대(Knowledge Enhancement), 발표, 출판 및 배포(Presentation, Publication and Dissemination), 이용자 경험(User Experience), 저장보관소 관리(Repository Management), 보존(Preservation)과 같이 8가지 요소로 나누어져 있고, 맥락 관리는 목표 및 사용 모델(Goal and Usage Modeling), 도메인 모델(Domain Modeling), 전거 관리(Authority Management)로 나누어져 있다.

### 3. DCC & U 생애주기 모델

DCC & U(Digital Curation Centre & Unit) 생애주기 모델은 DCC 생애주기 모델의 확장형이라고 할 수 있다. 〈그림 3〉에서 나타난 바와 같이 기존의 DCC 생애주기 프로세스 모델과 비교했을 때, DCC & U 생애주기 프로세스 모델은 데이터의 지식 증대와 관련하여 정보를 수정하는 활동인 큐레이트, 보존 및 지식 증대부분이 추가되었다. 또한, 정보에 관한 정보를 포함하는 기술 및 표현 정보 활동이 추가되며, 마지막으로 이용자 경험을 포함하는 연속 주기 동작이 추가되었다.<sup>31)</sup>

30) Panos Constantopoulos et al., "DCC&U: An Extended Digital Curation Lifecycle Model," *The International Journal of Digital Curation*, Vol.1, No.4(2009), p.40.

31) *Ibid.*, p.42.



〈그림 3〉 DCC & U 생애주기 모델<sup>32)</sup>

이상에서 살펴본 3개의 디지털 큐레이션 프로세스 모델의 구성요소들을 표로 정리하면 다음 〈표 1〉과 같다.

〈표 1〉 디지털 큐레이션 프로세스 모델 구성요소

	사전 계획	데이터 선정 및 입수			보존 활동			접근 및 이용			책임 및 관리
DCC	개념화	생성 및 접수	평가 및 선정	입수	보존	저장	변환	접근, 사용 및 재활용	-	-	-
DCU	-	평가	분류 색인 및 목록	입수	보존	-	-	발표, 출판 및 배포	이용자 경험	지식 증대	저장 보관소 관리
DCC&U	개념화	생성 및 접수	평가 및 선정	입수	보존	저장	변환	접근, 사용 및 재활용	이용자 경험	-	-

각각의 구성요소들을 크게 구분하면 사전 계획, 데이터 선정 및 입수, 보존 활동, 접근 및 이용, 책임 및 관리 5가지로 나눌 수 있다. 이 중 세 가지 모델의 구성요소들이 공통적으로 포함되어 있는 부분은 데이터 선정 및 입수, 보존 활동, 접근 및 이용 부분이라고 할 수 있다.

데이터의 선정 및 입수에는 수집할 데이터를 평가하고 선정하여 분류, 색인 및 목록과정 등을 거쳐 입수하는 과정이 포함된다. 보존 활동에는 보존, 저장, 변환의 과정이 포함되며, 접근 및 이용에는 접근, 사용 및 재활용, 저장된 데이터의 발표, 출판 및 배포, 이용자 경험, 지식 증대를 통한 접근이 포함된다.

32) Ibid.

## IV. 대학도서관의 디지털 큐레이션 프로세스 분석

앞의 3장에서 디지털 큐레이션 프로세스 모델을 살펴보았는데, 그 내용을 바탕으로 디지털 큐레이션 프로세스의 공통적인 구성요소를 크게 구분하면 데이터 선정 및 입수, 보존 활동, 접근 및 이용으로 간략화 할 수 있다. 이를 기반으로 대학도서관의 디지털 큐레이션 프로세스 및 보존 정책을 분석하기 위하여 2013년 Sheldon이 제안한 디지털 보존 모델의 19개 요소<sup>33)</sup>와 비교하여 재구성함으로써 본 논문에서 실시하는 대학도서관 디지털 큐레이션 프로세스 및 보존 정책의 분석 기준을 설정하였다.

디지털 큐레이션 프로세스 모델의 공통적인 구성 요소와 비교해 보면 콘텐츠 범위, 선정 및 평가, 신규자료 및 입수의 요소는 데이터 선정 및 입수에 포함될 수 있고, 저장, 복사 및 백업, 보안 관리, 협업, 메타데이터 및 기록문서는 보존 활동에 포함될 수 있다. 또한 접근 및 이용, 권리 및 규제 관리의 요소는 배포, 접근 및 사용에 포함될 수 있다. 나머지 요소들 중 앞선 디지털 큐레이션 프로세스 모델 구성요소에서 공통적인 요소로 추출되지는 못하였지만 전체적인 구성요소에서 나타난 부분이었던 책임 및 관리에 포함되는 요소들은 감사, 권한, 정책 및 전략 검토, 역할 및 책임, 담당자 훈련 및 교육, 보존 모델 및 전략, 보존 계획이 있다.

2013년 Sheldon이 제안한 디지털 보존 모델의 19개 구성 요소들을 디지털 큐레이션 프로세스 모델의 공통적인 구성 요소와 비교하여 내용상 밀접한 관계로 인해 분석 도중 중복의 우려가 있는 요소들을 통합하여 명칭을 변경하고, Sheldon이 제시한 구성요소에는 존재하지 않지만 분석 기준에 부합하는 요소를 추가하여 디지털 큐레이션 프로세스 분석 기준을 설정하면 다음 <표 2>와 같다.

4장에서는 <표 2>와 같은 분석 기준을 토대로 총 10개의 대학도서관을 대상으로 분석을 실시하였다. 기관의 선정은 Sheldon의 2013년 연구에서 나타난 16개의 대학도서관 가운데 성문화된 디지털 보존 정책을 수립하고 있는 8개의 대학도서관과 지역 내의 대학도서관을 모두 아우르는 덴마크의 대학도서관 디지털 보존 정책, 그리고 디지털 큐레이션, 디지털 저장 등 현재 활발하게 논의되고 있는 디지털 정보 및 환경에 관한 이슈 정보를 제공하는 Digital Scholarship에서 언급하고 있는 대학도서관 중 역시 성문화된 디지털 보존 정책을 수립하고 있는 Yale University Library를

33) 2013년 Sheldon은 아카이브, 도서관, 박물관의 디지털 보존 정책을 분석하면서 19개의 구성요소를 다음과 같이 추출하였다. 접근 및 이용(Access, Use), 신규자료 및 입수(Accessioning, Ingest), 감사, 심사(Audit), 서지(Bibliography), 협업(Collaboration), 콘텐츠 범위(Content Scope), 용어사전 및 전문용어(Glossary, Terminology), 권한(Mandates), 메타데이터 및 기록문서(Metadata and Documentation), 정책 및 전략 검토(Preservation Model, Strategy), 보존 모델 및 전략(Preservation Planning), 보존 계획(Policy, Strategy Review), 권리 및 규제 관리(Rights and Restriction Management), 역할 및 책임(Roles and Responsibilities), 보안 관리(Security Management), 선정 및 평가(Selection, Appraisal), 담당자 훈련 및 교육(Staff Training, Education), 저장, 복사 및 백업(Storage, Duplication, and Backup), 지속적 계획(Sustainability Planning)

〈표 2〉 디지털 큐레이션 프로세스 및 보존 정책 분석 요소 및 기준

분석 기준	분석 요소
데이터 선정 및 입수	선정 대상
	파일 포맷
	선정 및 평가
	입수
보존 활동	메타데이터
	저장, 복사 및 백업
	보안 및 위험 관리
	협업
접근 및 이용	접근 및 이용
	접근 권리 및 규제
책임 및 관리	감사
	역할 및 책임
	지속가능성
	보존 정책 및 전략 검토
	담당자 훈련 및 교육

추가하여 총 10개의 대학도서관의 디지털 큐레이션 및 보존 정책을 분석하였다.

분석 대상인 10개의 대학도서관은 다음 〈표 3〉과 같다.

〈표 3〉 디지털 큐레이션 프로세스 및 보존정책 분석 기관

번호	대학도서관
1	Johns Hopkins University Libraries
2	University of South Carolina Libraries
3	Boston University Libraries
4	University of Massachusetts Amherst Libraries
5	Dartmouth College Library
6	The National Library of Denmark and Copenhagen University Library
7	Statsbiblioteket State and University Library
8	University of Manchester Library
9	University of Utah Library
10	Yale University Library

### 1. 데이터 선정 및 입수

데이터 선정 및 입수는 데이터 선정 대상, 파일 포맷, 선정 및 평가, 입수의 요소가 포함되는 과정이다. 총 10개의 대학도서관 가운데 Boston University Libraries, The National Library of Denmark and Copenhagen University Library, Statsbiblioteket State and University Library를 제외한 7개의 대학도서관에서 선정 대상과 관련한 디지털 보존 정책을 성문화하고 있었으며, 선

정 대상은 대학 내에서 생산되거나 보관하고 있는 디지털 형태의 연구데이터를 중심으로 보존을 실시하였다. 뿐만 아니라 대학 내부에서 생산되지는 않았으나 대학에서 접근할 수 있도록 자격이 부여된 외부 기관의 디지털 데이터를 선정 대상으로 하였다.

10개 기관 중에서 디지털 보존정책에서 파일포맷을 제시하고 있는 기관은 5개 기관으로 이들 기관에서 수집하고 있는 파일포맷은 대부분 문서, 음성, 영상, 이미지포맷으로 이루어져 있었다. 파일 포맷과 관련한 디지털 보존 정책을 성문화하여 제시하고 있는 대학도서관과 그 내용을 정리 하면 다음 <표 4>와 같다.

<표 4> 대학도서관별 디지털 보존 파일포맷

	파일 포맷
Johns Hopkins University Libraries	Text : ASCII, UTF-8 Unicode
	Audio : AIFF, Wave, MP2
	Video : MPEG, MPEG-2, AVI, MPEG-7
	Image : Gif, JPEG, JPEG2000, JFIF, TIFF, ITU-T.6
Boston University Libraries	Text : HTML, doc, ppt, xls, PDF, ps, rtf, txt, SGML, XML, DTD
	Audio : MPEG, Real Audio, Wave, WMA
	Video : AVI, MPEG-1, MPEG-2, MPEG-4, Quicktime, WMV
	Image : BMP, GIF, JPEG, JPEG2000, PNG, Photo CD, Photoshop, TIFF
Dartmouth College Library	Text : ASCII, UTF-8, Unicode
	Audio/Video : MPEG, AVI, MOV, AAC, WAV
	Image : TIFF, JPEG, GIF, JPEG2000
	Datasets : XML, XLS
University of Utah Library	Text : PDF/A-1 ISO 19005-1
	Audio : FLAC
	Video : Motion JPEG 2000
	Image : TIFF
	Web : XML (includes XSD/ XSL/XHTML)
Yale University Library	Text : ASCII, UTF-8 Unicode
	Audio : AIFF, Wave, MP2
	Video : MPEG, MPEG-2, AVI
	Image : Gif, JPEG, JPEG2000, TIFF

<표 4>에 나타난 결과를 보면 파일 포맷은 대부분 문서, 음성, 영상, 이미지로 나누어 저장하고 있음을 알 수 있고, 문서는 ASCII와 PDF, 음성은 WAV, 영상은 MPEG, 이미지는 TIFF 포맷이 가장 많이 저장에 이용됨을 알 수 있다.

다음으로 디지털 보존을 실시할 데이터의 선정 및 평가와 입수에 관한 대학도서관별 정책을 살펴보면 데이터의 선정 및 평가는 University of South Carolina Libraries, University of

Massachusetts Amherst Libraries, Dartmouth College Library, University of Utah Library에서 나타나고 있었으며, 그 내용은 University of South Carolina Libraries와 Dartmouth College Library의 경우는 기본적으로 도서관의 수집 정책을 바탕으로 디지털 형태의 자료를 포함하도록 명시되어 있었다. University of Massachusetts Amherst Libraries와 University of Utah Library의 정책에는 교육 과정 및 연구와의 관련성, 기존 소장 자료들과의 관련성, 데이터의 가용성, 소장 하드웨어 및 플랫폼에서의 사용가능성, 구입 비용 및 예상 이용자, 장기적 보존의 필요성 등을 선정 및 평가 기준으로 제시하였다.

입수에 관한 정책은 Dartmouth College Library와 StatsBiblioteket State and University Library에서 나타나고 있으며, Dartmouth College Library는 담당자 확인, 파일 형식 확인, 내용 검증, 메타데이터 생성, 데이터 변환의 과정을 거쳐 입수하도록 하고 있었다. 덴마크 대학도서관의 입수 과정은 생산자 확인, 데이터 변환의 필요성 확인, 데이터 변환, 저장 파일 포맷 선정, 포맷 변환의 과정을 거친다.

## 2. 보존 활동

보존 활동에는 메타데이터, 저장, 복사 및 백업, 보안 및 위험 관리, 협업의 요소가 포함된다. 먼저 메타데이터에 관한 정책을 살펴보면 메타데이터는 대학도서관마다 정책에 따라 다른 메타데이터를 사용하도록 되어있다. 분석 결과에서는 Johns Hopkins University Libraries와 Dartmouth College Library에서 전자정보의 장기 보존을 위한 가장 종합적이고 포괄적인 메타데이터 요소 스키마로 받아들여지고 있는 PREMIS(Preservation Metadata Implementation Strategies)를 사용하고 있었으며, University of South Carolina Libraries와 University of Manchester Library에서는 도서관의 표준 메타데이터인 더블린코어(Dublin Core)를 사용하고 있었다. Yale University Library의 경우 디지털 도서관 환경에 적합한 기술용 메타데이터인 MODS(Metadata Object Description Schema)를 사용하고 있는 것으로 나타났다.

저장, 복사 및 백업과 관련해서는 Johns Hopkins University Libraries와 Dartmouth College Library, Yale University Library에서 성문화된 정책으로 나타나고 있었는데, 공통적으로는 데이터 손상 및 소실의 위험을 줄이기 위해 물리적으로 거리가 떨어진 보존 시스템에 데이터 사본을 저장하는 내용을 기본으로 하는 양상을 보이고 있었다. Johns Hopkins University Libraries는 백업 프로그램을 문서화하고, 타 기관과의 협력을 통한 원격 백업을 실시하도록 하고 있었으며, Dartmouth College Library의 경우에는 하드웨어, 소프트웨어, 네트워크 액세스와 같은 기술 인프라를 구축할 것을 명시하고 있었다. Yale University Library의 경우는 분산 이기종 시스템을 운영하도록 하고 있으며, 하드웨어 마이그레이션 서비스를 실시하도록 하고 있었다.

보존 활동에서 보안 및 위험 관리는 10개의 대학도서관 가운데 6개인 Johns Hopkins University Libraries, University of South Carolina Libraries, University of Massachusetts Amherst Libraries, The National Library of Denmark and Copenhagen University Library, StatsBiblioteket State and University Library, University of Utah Library에서 성문화된 정책으로 나타나고 있다. 6개 대학도서관 모두 디지털 데이터의 위험 요소를 단순한 기술적인 환경변화 뿐만 아니라 재정적 위험, 시스템의 위험, 물리적 공간의 위험, 이용자 및 담당자 접근 시 조작의 위험으로 나누어 제시하고 있었으며, 이를 위해 위험 요소에 대한 평가 과정을 실시하고 장기적 재정적 지원의 필요성을 명시하고 있다.

보존 활동에서 협업과 관련된 부분은 Johns Hopkins University Libraries, StatsBiblioteket State and University Library, Yale University Library를 제외한 7개의 대학도서관에서 나타나고 있는데, 협업은 크게 대학 내부의 기관과 협력하는 방법과 대학 외부의 기관과 협력하는 방법으로 나누어진 다. University of South Carolina Libraries와 University of Massachusetts Amherst Libraries의 경우는 대학 내부의 대학도서관, 대학 당국, 대학의 정보 기술 센터와 협력하여 전문 기술 부분에 대한 도움을 제공받을 수 있도록 정책적으로 제시하고 있었으며, Boston University Libraries의 경우는 Portico, LOCKSS와 협업하고 있음을 명시하고 있다. Dartmouth College Library의 경우는 대학 외부의 기관과의 협력을 통해 데이터의 공동 활용과 보존에의 공동 노력을 진행하고 있었으며, The National Library of Denmark and Copenhagen University Library는 덴마크에서 보존을 담당하고 있는 기타 기관들과 협업하고 있다. University of Manchester Library는 세계적인 기구들을 대상으로 협업을 진행하도록 하고 있었으며, 마지막으로 University of Utah Library에서는 지역적 문화 유산 기관이나 국가 문화 유산 기관 등과의 협업을 진행하고 있는 것으로 나타났다.

### 3. 접근 및 이용

접근 및 이용에 관한 정책은 10개의 대학도서관 가운데 Johns Hopkins University Libraries, Boston University Libraries, University of Manchester Library를 제외한 7개의 대학도서관에서 나타나고 있었으며, 개인 정보 보호법, 저작권법 등 법률에 관련한 부분과 외부 기관 저장소에 접근하는 경우에는 계약을 통한 인증 등 규약 및 법칙에 관련된 부분이 공통적으로 나타나고 있다. University of South Carolina Libraries의 경우에는 관련 법규, 규정, 허가, 계약 등으로 정의되어 있는 접근 제한을 실시하고 있었으며, University of Massachusetts Amherst Libraries와 Dartmouth College Library는 지적 재산권, 법적 요구 사항, 개인 정보 보호법, 프로젝트의 임무 등에 따라 접근 권한을 설정하고 있다. The National Library of Denmark and Copenhagen University Library의 경우는 규제, 제한, 등록의 3가지 형태로 나누어 접근을 허용하고 있는데, 지적 재산권이나 개인 정보 보호법 등으로 규제되는 데이터와, 타 기관에서의 접근이나 이용을

실시할 경우 제한을 두고 있으며, 대학 및 대학도서관에 소속되어 있거나 협력 기관일 경우 등록되어 이용할 수 있도록 하고 있다. StatsBiblioteket State and University Library, University of Utah Library, Yale University Library의 경우에는 법적 권리 및 기술적 문제를 확인하여 도서관에 소속되어 있거나, 후원 기관, 협력 기관 등에서 접근 및 이용할 수 있도록 하고 있다.

접근 권리 및 규제에 관한 정책은 The National Library of Denmark and Copenhagen University Library에서만 명시하고 있는데, 관련 법규, 규정, 법적 요구 사항, 개인 정보 보호법에 따라 결정하도록 되어 있다.

#### 4. 책임 및 관리

책임 및 관리 부분은 감사, 역할 및 책임, 지속가능성, 보존 정책 및 전략 검토, 담당자 훈련 및 교육의 요소로 나누어 분석할 수 있는데, 먼저 감사의 부분에 관하여 정책적으로 명확하게 명시되어 있는 대학도서관은 Johns Hopkins University Libraries, Dartmouth College Library, The National Library of Denmark and Copenhagen University Library, StatsBiblioteket State and University Library의 4개이다. Johns Hopkins University Libraries, The National Library of Denmark and Copenhagen University Library, StatsBiblioteket State and University Library는 기술적 부분, 정책적 부분, 관리 부분, 재정적 부분, 인력 부분으로 나누어 감사를 실시하고 있었으며, Dartmouth College Library에서는 TRAC(Trustworthy Repositories Audit & Certification) 기준에 맞추어 감사를 실시하고 있었다.

역할 및 책임에 관련해서는 The National Library of Denmark and Copenhagen University Library와 University of Manchester Library를 제외한 8개의 대학도서관에서 정책적으로 제시하고 있다. Johns Hopkins University Libraries, Boston University Libraries, Dartmouth College Library의 경우에는 대학도서관의 단일 기관을 중심으로 디지털 데이터의 보존 노력에 대한 역할 및 책임을 강조하고 있었다. 반면에 University of South Carolina Libraries, University of Massachusetts Amherst Libraries, StatsBiblioteket State and University Library, University of Utah Library, Yale University Library에서는 보존 활동과 관련된 기타 내부 협력 기관들의 역할 및 책임까지 나타내면서 대학 내의 협력 기관들과의 공동 책임을 강조하고 있는 것으로 나타났다. 특히 기술(technical)과 같은 전문적인 부분에서는 대학 내의 정보 기술 센터 등과 역할 및 책임을 공유할 수 있도록 하는 내용을 명시하고 있었다.

책임 및 관리에 관한 분석에서 지속가능성, 보존 정책 및 전략, 담당자 훈련 및 교육 요소에 대한 정책은 많은 대학도서관에서 실시하고 있지 않는 것으로 나타났다. 대학도서관별 지속가능성을 먼저 살펴보면 Johns Hopkins University Libraries, The National Library of Denmark and

Copenhagen University Library, StatsBiblioteket State and University Library에서 나타나고 있었으며, Johns Hopkins University Libraries는 기술 인프라 평가, 인력 부분 관리, 재정적 부분 관리, 관리 감독 부분으로 나누어 지속가능성을 평가하고 있었다. The National Library of Denmark and Copenhagen University Library는 특히 재정적 지원 부분에 대한 지속가능성을 강조하고 있었으며, StatsBiblioteket State and University Library에서는 보존 계획을 바탕으로 디지털 데이터에 대한 정책, 위험 관리, 재정적 지원 부분에서 지속가능성을 평가하고 있는 것으로 나타났다.

보존 정책 및 전략 검토는 4개의 대학도서관에서 성문화된 정책으로 나타나고 있었는데, University of South Carolina Libraries는 각 학기가 시작되는 시기 또는 필요한 경우에는 더욱 자주 검토될 수 있도록 하고 있고, Dartmouth College Library는 3년 한 번 이상 정책 및 전략 검토를 실시하도록 하고 있다. The National Library of Denmark and Copenhagen University Library는 2년에 한번, Yale University Library는 매년 검토하도록 정책으로 제시되고 있었다.

마지막으로 담당자 훈련 및 교육 부분은 Dartmouth College Library와 The National Library of Denmark and Copenhagen University Library에서 나타나고 있는데 Dartmouth College Library에서는 디지털 보존과 관련하여 담당하고 있는 직원들을 계속적으로 교육시킬 것을 강조하고 있었으며, The National Library of Denmark and Copenhagen University Library에서는 내부 업무 수행 및 아웃소싱을 담당하는 충분한 인력을 배치할 수 있도록 하고, 지속적이고 정기적인 훈련 및 교육이 실시될 수 있도록 하고 있었다.

## 5. 종합분석

이상에서 국외 10개 대학도서관의 디지털 큐레이션 프로세스 및 보존 정책을 데이터 선정 및 입수, 보존 활동, 접근 및 이용으로 구분하여 분석하였다. 그 결과를 간단하게 표로 정리하면 다음 <표 5>와 같다.

표 5)에 나타난 결과를 보면, 데이터 선정 및 입수 과정에서는 선정 대상에 대한 정책이 7개 대학도서관, 파일 포맷에 대한 정책이 5개 대학도서관, 선정 및 평가에 대한 정책이 4개 대학도서관, 입수에 대한 정책이 2개 대학도서관으로 나타났다.

선정 대상은 대부분 대학 내에서 생산되거나 보관하고 있는 디지털 형태의 연구데이터를 중심으로 대학에서 접근할 수 있도록 자격이 부여된 외부 기관의 디지털 데이터가 주를 이루었고, 파일 포맷은 문서, 음성, 영상, 이미지로 나누어 문서는 ASCII와 PDF, 음성은 WAV, 영상은 MPEG, 이미지는 TIFF 포맷이 가장 많이 저장에 이용되었다. 선정 및 평가에서는 공통적으로 이용 가능성과 장기적 보존의 필요성, 기존 대학도서관의 소장 데이터와의 관련성 등을 평가하였고, 입수에서는 생산자 확인, 데이터 변환, 포맷 변환을 거쳐 메타데이터를 생성하고 보존 저장소로 보내졌다.

<표 5> 대학도서관의 디지털 큐레이션 프로세스 및 보존 정책

		Johns Hopkins University Libraries	University of South Carolina Libraries	Boston University Libraries	University of Massachusetts Amherst Libraries	Dartmouth College Library
국가		미국	미국	미국	미국	미국
정책 생성 년도		2008	2010	2010	2011	2011
정책 갱신 년도		2008	2010	2010	2011	2011
데이터 선정 및 입수	선정 대상	○	○		○	○
	파일 포맷	○		○		○
	선정 및 평가		○		○	○
	입수					○
보존 활동	메타데이터	○	○			○
	저장, 복사 및 백업	○				○
	보안 및 위험 관리	○	○		○	
	협업		○	○	○	○
접근 및 이용	접근 및 이용		○		○	○
	접근 권리 및 규제					
책임 및 관리	감사	○				○
	역할 및 책임	○	○	○	○	○
	지속가능성	○				
	보존 정책 및 전략 검토		○			○
	담당자 훈련 및 교육					○
		The National Library of Denmark and Copenhagen University Library	Statsbiblioteket State and University Library	University of Manchester Library	University of Utah Library	Yale University Library
국가		덴마크	덴마크	영국	미국	미국
정책 생성 년도		2009	2012	2012	2010	2005
정책 갱신 년도		2012	2012	2012	2012	2007
데이터 선정 및 입수	선정 대상			○	○	○
	파일 포맷				○	○
	선정 및 평가				○	
	입수		○			
보존 활동	메타데이터			○		○
	저장, 복사 및 백업					○
	보안 및 위험 관리	○	○		○	
	협업	○		○	○	
접근 및 이용	접근 및 이용	○	○		○	○
	접근 권리 및 규제	○				
책임 및 관리	감사	○	○			
	역할 및 책임		○		○	○
	지속가능성	○	○			
	보존 정책 및 전략 검토	○				○
	담당자 훈련 및 교육	○				

보존 활동 부문에 있어서는 메타데이터에 대한 정책이 5개 대학도서관, 저장, 복사 및 백업에 대한 정책이 3개 대학도서관, 보안 및 위험 관리에 대한 정책이 6개 대학도서관, 협업에 대한 정책이 7개 대학도서관에서 성문화되어 나타났다.

메타데이터는 대학도서관마다 정책에 따라 다른 메타데이터를 사용하도록 되어있었는데, 분석에서는 PREMIS와 더블링크어를 사용하는 대학도서관이 각각 2개씩 나타났다. 저장, 복사 및 백업 정책에서는 데이터 손상 및 소실의 위험을 줄이기 위해 물리적으로 거리가 떨어진 보존 시스템에 데이터 사본을 저장하는 내용을 기본으로 하는 양상을 보이고 있었다. 보안 및 위험 관리에서 위험 요소들은 단순한 기술적인 환경변화 뿐만 아니라 재정적 위험, 시스템의 위험, 물리적 공간의 위험, 이용자 및 담당자 접근 시 조작의 위험 등으로 나누어 나타났으며, 협업은 대학 내부의 기관과 협력하는 방법과 대학 외부의 기관과 협력하는 방법으로 나누어지는 양상을 보였다.

접근 및 이용 부분에서는 접근 및 이용에 관한 정책을 수립하고 있는 대학도서관이 7개, 접근 권리 및 규제에 관한 정책을 수립하고 있는 대학도서관이 1개로, 접근 권리 및 규제에 관한 정책이 가장 저조한 형태로 나타났다. 데이터에 대한 접근, 이용, 규제는 모두 개인 정보 보호법, 저작권법 등 법률과 관련한 부분과 외부 기관 저장소에 접근하는 경우에는 계약을 통한 인증 등 규약 및 법칙에 관련된 부분이 공통적으로 나타났다.

마지막으로 책임 및 관리 부분에서는 감사에 관한 정책이 4개, 역할 및 책임에 관한 정책이 8개, 지속가능성에 대한 정책이 3개, 보존 정책 및 전략 검토에 관한 정책이 4개, 담당자 훈련 및 교육 정책이 2개로 나타났다.

감사는 기술적 부분, 정책적 부분, 관리 부분, 재정적 부분, 인력 부분 등으로 나누어 감사를 실시하고, 역할 및 책임은 특히 보존의 측면에 있어서는 대학도서관 단독의 역할 및 책임이 아니라 대학 내의 기타 협력 기관들과의 공동 책임을 강조하는 모습을 보였다. 지속가능성 역시 감사와 비슷하게 기술, 정책, 관리, 재정, 인력 부분 등으로 나누어서 평가하도록 하고, 보존 정책 및 전략 검토는 최신의 흐름을 반영하여 즉각적인 피드백이 일어날 수 있도록 하는 형태로 구성되어 있었다. 마지막으로 담당자 훈련 및 교육 부분은 보존을 담당하는 담당 인력에 대한 교육 정책으로써 특히 지속적이고 정기적인 훈련 및 교육을 강조하고 있었다.

이상에서 국외 대학도서관 10곳의 디지털 큐레이션 프로세스 및 보존 정책을 분석하였다. 그 결과를 토대로 국내 대학도서관의 디지털 큐레이션 서비스 도입을 위해데이터 선정 및 입수, 보존 활동, 접근 및 이용구분해서 방안을 제시하면 다음과 같다.

선정 데이터는 대학 내에서 생산되는 연구데이터로 한정하고 보존 파일 포맷은 분석 결과에서 나타난 바와 같이 문서의 경우 PDF, 음성의 경우 WAV, 영상의 경우 MPEG, 이미지의 경우 TIFF 포맷을 적용하는 것이 바람직할 것으로 보인다. 메타데이터 역시 대학도서관에서 사용할

것임을 감안할 때 더블린코어 혹은 PREMIS가 적합할 것이며, 디지털 데이터가 갖는 손상 및 손실의 위험을 줄이기 위해 주요 저장보관소 외에도 공간적으로 거리가 있는 곳에 백업 저장보관소를 구축하는 것이 바람직할 것이다. 보존 정책 및 전략에 관해서는 기술적 변화나 최신의 연구 동향에 맞추기 위해 1년 이하의 기간 내에 정책 검토를 실시하도록 하고, 담당자 훈련 및 교육 역시 1, 2년 간격을 두고 지속적으로 실시해 나가야 할 것이다.

디지털 데이터를 대상으로 하는 큐레이션에 있어서 보존은 매우 중요한 부분이다. 손상되기 쉬운 디지털 데이터 자체의 위험 뿐만 아니라 급변하는 디지털 환경에서 하드웨어 및 소프트웨어의 변화에 따라 보존 및 접근에 대한 위험은 항상 경계해야 하는 부분이다. 그러나 이러한 보존이나 저장소, 하드웨어 및 소프트웨어, 시스템 등의 관리는 전문적인 기술력이 요구된다. 그렇기 때문에 대학도서관은 대학도서관 내의 시스템관리 부서 혹은 대학 자체의 정보 기술 센터 등과 협력 관계를 구축하여 도움을 받을 필요가 있다. 뿐만 아니라 디지털 큐레이션 프로세스를 진행하고 디지털 보존을 위한 기술적 노력인 노력을 기울이기 위해서는 지속적인 재정적 지원이 뒷받침되어야 하는데 이는 대학 당국에서 지원해야 하는 부분이라 할 수 있다.

한편, 대학도서관에서 디지털 큐레이션 서비스를 도입하기 위해서는 대학 내의 협력 관계의 구축이 매우 중요한 부분이 될 것이다. 대학에서는 디지털 데이터의 보존 및 큐레이션을 위해 지속적이고 안정된 재정적 지원을 실시하고, 정보 기술 센터는 보존활동에 있어서 필요한 기술적인 부분을 담당하여 보존 및 저장소를 관리할 필요가 있을 것이다. 도서관은 수많은 자료들 중 정제된 양질의 정보를 선별적으로 수집, 관리, 제공하는 기관이다. 기관의 본 목적에 맞게 큐레이션 프로세스에서도 대학도서관에서는 질적인 측면에서 수집할 데이터를 선정하여 입수하고, 이용자들이 필요한 데이터에 접근하여 이용할 수 있도록 서비스를 제공하는 측면에 초점을 맞추어 강화하는 것이 바람직할 것이다.

## V. 결 론

연구데이터는 연구자들의 연구 또는 연구수행과정 중에 산출되는 데이터로, 이러한 연구데이터는 연구결과에 대한 핵심정보를 담고 있기 때문에 연구결과를 증명할 수 있는 중요한 자료라고 할 수 있다. 본 논문에서는 이러한 연구데이터의 효율적인 활용을 위해 디지털 큐레이션 프로세스 및 보존 정책을 4가지 부분, 15개의 요소로 국외의 10개 대학도서관을 대상으로 분석하였다.

총 10개의 대학도서관의 디지털 큐레이션 프로세스 및 보존 정책에 대해 15개의 분석 기준을 설정하여 분석한 결과를 간단하게 정리하면 다음과 같다.

첫째, 데이터 선정 및 접수에서 수집 데이터는 대학 내에서 생산된 디지털 형태의 연구데이터나

외부 기관에서 생성된 디지털 데이터지만 접근 및 이용에 대한 권한을 갖고 있는 데이터가 대부분이며, 보존을 실시할 때의 포맷은 문서의 경우 ASCII, PDF, 음성의 경우 WAV, 영상의 경우 MPEG, 이미지의 경우 TIFF 포맷을 가장 많이 사용하였다. 평가를 통해 입수하는 과정에서는 공통적으로 이용가능성과 장기적 보존의 필요성, 기존 소장 데이터와의 관련성을 기준으로 평가하고 데이터 생산자, 데이터 변환, 포맷 변환을 확인하고 저장 보관소로 송부하는 과정을 거치는 것으로 나타났다.

둘째, 보존 활동에 있어서 메타데이터는 도서관마다 선택적으로 사용하지만 대학도서관에서는 표준 메타데이터인 더블링크어와 PREMIS를 일반적으로 사용하고 있으며, 디지털 데이터의 손상 및 손실과 같은 위험을 방지하기 위해 사본을 제작해 물리적으로 거리가 먼 시스템에 백업하고 있었다. 협업은 많은 대학도서관의 정책에서 나타나고 있는 요소인데 대학 내부에서는 대학도서관과 대학 당국, 대학의 정보 기술 센터와 협력함으로써 재정적 역할과 기술적 역할을 담당하도록 하고, 외부 협업을 통해서도 데이터 접근의 확장과 보존에 대한 공동 노력을 펼칠 수 있도록 하는 것으로 나타났다.

셋째, 접근 및 이용에서는 저작권법이나 개인정보보호법과 같은 법률, 외부 기관과의 계약 접근 허가 등 법률 및 규약에 기초하여 접근, 규제 등을 결정할 수 있도록 하고 있었다.

넷째, 책임 및 관리 차원에서는 대부분의 요소에서 관리적, 재정적, 기술적, 물리적, 인적 부분 등으로 나누어 각각의 감사나 지속가능성 등을 검토하도록 하고, 역할 및 책임에서도 대학도서관 단독적 역할이나 책임이 아닌 교내 기관 또는 대학 외부 기관들과의 공동 책임임을 명시하고 있는 것으로 나타났다.

## 참고문헌

- 국립중앙도서관 도서관연구소(KRILI). “도서관용어해설.” 도서관연구소 웹진 32(2009).
- 김선태 등. “과학데이터 보존 및 활용모델에 관한 연구.” 한국비블리아학회지, 제21권, 제4호(2010), pp.81-93.
- 김은정. 연구데이터 수집에 영향을 미치는 요인 분석. 박사학위논문, 중앙대학교 대학원 기록관리학과, 2012.
- 김은정, 남태우. “연구데이터 수집에 영향을 미치는 요인 분석.” 정보관리학회지, 제29권, 제2호(2012), pp.27-44.
- 서울대학교, 서울대학교 연구윤리지침 <[http://nrf.re.kr/nrf\\_tot/cms/include/download\\_general.jsp?file=d\\_0204\\_63513\\_0\\_98470.pdf](http://nrf.re.kr/nrf_tot/cms/include/download_general.jsp?file=d_0204_63513_0_98470.pdf)> [인용 2014. 5. 7].

- 설문원 등. 국가 디지털 아카이빙 체제 구축에 관한 연구. 서울:한국과학기술정보연구원, 2005.
- 안영희, 박옥화. “대학도서관 서비스의 디지털 큐레이션 전략.” 한국도서관·정보학회지, 제40권, 제4호(2009), pp.311-326.
- 안영희, 박옥화. “디지털 큐레이션 정책을 위한 프레임워크 개발.” 한국도서관·정보학회지, 제41권, 제1호(2010), pp.167-186.
- 현문수. “디지털 연구데이터 장기보존의 편익에 대한 연구.” 한국기록관리학회지, 제11권, 제1호(2011), pp.161-181.
- Abbott, D. “What is Digital Curation?” *DCC Briefing Papers: Introduction to Curation*, Edinburgh: Digital Curation Centre, 2008.
- Boston University Libraries: Digital Initiatives & Open Access - Digital Preservation Policy <<http://www.bu.edu/dioa/openbu/boston-university-libraries-digital-preservation-policy/>> [cited 2014. 3. 10].
- Dartmouth College Library - Digital Preservation Policy <<http://www.dartmouth.edu/~library/digital/about/policies/preservation.html?mswitch-redirect=classic>> [cited 2014. 3. 10].
- Digital Curation Centre <<http://www.dcc.ac.uk/>> [cited 2014. 5. 7].
- EMC & IDC. IDC Digital Universe Study: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East <[www.emc.com/collateral/analyst-reports/idc-digital-universe-united-states.pdf](http://www.emc.com/collateral/analyst-reports/idc-digital-universe-united-states.pdf)> [cited 2014. 5. 7].
- Heidorn, P. Bryan. “The Emerging Role of Libraries in Data Curation and E-science.” *Journal of Library Administration*, Vol.51, No.7(2011), pp.662-672.
- John Hopkins Sheridan Libraries - JScholarship Digital Preservation Policy <<http://old.library.jhu.edu/collections/institutionalrepository/irpreservationpolicy.html>> [cited 2014. 3. 10].
- Jones, Sarah. A Report on the Range of Policies for and Related to Digital Curation <<http://www.emc.com/collateral/analyst-reports/idc-digital-universe-united-states.pdf>> [cited 2014. 5. 7].
- Jubb, Michael. “UK Research Funders’ Policies for the Management of Information Outouts.” *International Journal of Digital Curation*, Vol.2, No.1(2007), pp.29-48.
- Madeline Sheldon. Analysis of Current Digital Preservation Policies: Archives, Libraries, and Museums. Junior Fellow with NDIIPP. 2013.
- National Academy of Sciences. Committee on Ensuring the Utility and Integrity of Research Data in a Digital Age, Ensuring the integrity, accessibility, and stewardship of research data in the digital age. Washington, D.C.: National Academy Press, 2009.
- OECD. OECD principles and guidelines for access to research data from public funding

- 〈<http://www.oecd.org/science/sci-tech/38500813.pdf>〉 [cited 2014. 5. 7].
- Panos Constantopoulos, Costis Dallas, Ion Androutopoulos, Stavros Angelis, Antonios Deligiannakis, Dimitris Gavrilis, Yannis Kotidis, Christos Papatheodorou, "DCC&U: An Extended Digital Curation Lifecycle Model." *The International Journal of Digital Curation*, Vol.1, No.4(2009), pp.34-45.
- Ray, Joyce. *Managing the Digital World: the Role of Digital Curation*, UK:Edinburgh, 2008.
- Sarah Higgins. "The DCC Curation Lifecycle Model." *The International Journal of Digital Curation*, Vol.1, No.3(2008), pp.134-140.
- Tyler O. Walters. "Data Curation Program Development in U.S. Universities: The Georgia Institute of Technology Example." *The International Journal of Digital Curation*, Vol.3, No.4(2009), pp.83-92.
- UMass Amherst Libraries, University of Massachusetts Amherst Libraries Collection Development Policy 〈<http://www.library.umass.edu/about-the-libraries/library-policies-procedures-and-guidelines/collection-development-policy/>〉 [cited 2014. 3. 10].
- University of Illinois at Urbana-Champaign, Data Curation Education Program 〈<http://www.lis.illinois.edu/research/projects/data-curation-education-program>〉 [cited 2014. 6. 8].
- University of Manchester Library - policies 〈<http://www.library.manchester.ac.uk/aboutus/policies/>〉 [cited 2014. 3. 10].
- University of Massachusetts Amherst Libraries - Collection Development Policy  
〈<http://www.library.umass.edu/about-the-libraries/library-policies-procedures-and-guidelines/collection-development-policy/>〉 [cited 2014. 3. 10].
- University of Michigan, Master of Science in Information, Preservation of Information specialization  
〈<http://www.si.umich.edu/academics/msi/preservation-information-pi>〉 [cited 2014. 6. 8].
- University of North Carolina at Chapel Hill, Certificate in Digital Curation  
〈[http://sils.unc.edu/programs/certificates/digital\\_curation](http://sils.unc.edu/programs/certificates/digital_curation)〉 [cited 2014. 6. 8].
- University of South Carolina Libraries - About Digital Collections  
〈<http://library.sc.edu/digital/about.html>〉 [cited 2014. 3. 10].
- University of South Carolina Libraries - Collection Development: General Policies  
〈<http://library.sc.edu/techserv/cmgenpol.html>〉 [cited 2014. 3. 10].
- Uribe, Lusi Martines, and Macdonald, Stuart. "A new role for academic librarians: data curation." *Profesinoal de la Informacion*, Vol.17, No.3(2008), pp.273-280.
- Yakel, Elizabeth. "Archives and Manuscripts Digital Curation, OCLC Systems & Services."

*International digital library perspectives*, Vol.23, No.4(2007), pp.335-340.

국한문 참고문헌의 영어표기

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

Ahn, Young-Hee, & Park, Ok-Wha. "Digital Curation Strategy for University Library Service."

*The Korea Journal of Library and Information Science*, Vol.40, No.4(2009), pp.311-326.

Ahn, Young-Hee, & Park, Ok-Wha. "Development of a Framework for Digital Curation Policy."

*The Korea Journal of Library and Information Science*, Vol.41, No.1(2010), pp.167-186.

Hyun, Moon-Soo. "A Study on Benefits of Digital Preservation of Research Data." *The Korea*

*Journal of Information Management*, Vol.11, No.1(2011), pp.161-181.

Kim, Eun-Jeong. Factor Analysis of Effects on Research Data Collection. The Department of Record Management. The Graduate School of Chung-ang University. 2012.

Kim, Eun-Jeong, & Nam, Tae-Woo. "Factor Analysis of Effects on Research Data Collection."

*The Korea Journal of Information Management*, Vol.29, No.2(2012), pp.27-44.

Kim, Sun-Tae. et al. "A Study on a Model for Using and Preserving Scientific Data." *The Korean Journal of Biblia*, Vol.21, No.4(2010), pp.81-93.

KRILII(Korea Research Institute for Library and Information). "Libraries Glossary of Terms."

*Research Institute for Library Webzine*. 32(2009).

Seol, Moon-Won, Lee, So-Yeon, Lee, Yong-Hoon & Lee, Mi-Hwa. A Study on Developing National Digital Archiving Strategies. Seoul:Korea Institute of Science and Technology Information, 2005.

Seoul National University. Seoul National University Research Ethics Guidelines. 2010.