

디지털 큐레이션 정책을 위한 프레임워크 개발

Development of a Framework for Digital Curation Policy

안 영 희(Young-Hee Ahn)*

박 옥 화(Ok-Wha Park)**

〈 목 차 〉

- | | |
|----------------------|-----------------------------|
| I. 서 론 | 4. 디지털 큐레이션 정책 사례 |
| 1. 연구의 필요성 | III. 디지털 큐레이션 정책 프레임워크 개발 |
| 2. 연구 방법 | 1. 디지털 큐레이션 정책 프레임워크 구조 |
| 3. 선행 연구 | 2. 생명주기 기반 큐레이션 정책 프레임워크 개발 |
| II. 디지털 큐레이션 정책 및 사례 | IV. 결론 및 제언 |
| 1. 리포지터리와 데이터 센터 정책 | |
| 2. 데이터 관리와 공유 정책 | |
| 3. 기관 큐레이션 정책 | |

초 록

디지털 정보자원의 장기보존 및 접근, 공유, 재사용의 필수조건으로 디지털 큐레이션 정책이 수립되어야 한다. 이를 위하여 영국 디지털 큐레이션 센터 및 8개 연구재단의 디지털 큐레이션 정책과 지원 서비스 사례를 통하여 디지털 정보자원의 생명주기 단계에 따른 큐레이션 정책 프레임워크를 개발하였다. 본고에서 제안한 프레임워크는 국가차원이나 개별 조직차원에서 포괄적인 디지털 정보자원의 큐레이션 정책을 수립하기 위한 기본 틀을 제공할 것으로 기대된다.

키워드: 디지털 정보자원, 디지털 큐레이션, 생명주기, 장기보존 정책, 큐레이션 정책

ABSTRACT

This study aims at developing a policy framework for a digital curation based on the access, co-ownership, reuse, and long-term preservation of digital information resources. To accomplish this, our study analysed a similar policy and support services of a main UK research institution and the DCC. It selected the main stages of the DCC lifecycle and developed a policy framework accordingly. Such a framework could serve as the basis for both state and individual institutions, and for developing standards and guidelines for a digital curation of digital information resources.

Keywords: Curation Policy, Digital Curation, Digital Resources, Lifecycle, Long-term Preservation Policy

* 백석대학교 학술정보관 사서(yha@bu.ac.kr)(제1저자)

** 충남대학교 사회과학대학 문헌정보학과 교수(owpark@cnu.ac.kr)(공동저자)

• 접수일: 2010년 2월 26일 • 최초심사일: 2010년 3월 6일 • 최종심사일: 2010년 3월 26일

I. 서론

1. 연구의 필요성

정보통신기술의 발전 및 인터넷 환경의 변화로 디지털 정보자원의 생성이 기하급수적으로 늘어나고, 모든 학문분야에 걸쳐 디지털 정보자원의 제공, 활용 및 서비스가 다각적으로 이루어지고 있다. 한편 과학 분야뿐만 아니라 여러 영역의 디지털 정보자원 중 연구데이터에 대한 공유, 접근 및 재사용에 대한 관심이 모아지면서 2009년 11월 미국학술원에서는 연구데이터를 효율적으로 관리하고 활용할 수 있도록 3대 원칙(완전성, 접근성, 관리성)과 11개 권고안을 제시하였다.¹⁾ 이는 연구결과물이 체계적으로 관리되어 연구데이터에 대한 접근과 공유가 원활하도록 큐레이션이 효과적으로 이루어지게 하려는 권고안이다.

큐레이션은 연구데이터 관리와 공유를 통해 연구자들로 하여금 이미 발표되었거나 출간된 자료를 입증할 수 있게 하며, 새로운 분석 방법에 보다 쉽게 접근할 수 있게 해준다. 또한 기존의 연구자들이 밝혀낸 연구 자료에서 예상치 못한 질문을 통하여 데이터를 변경할 수도 있으며, 자료들을 통합하고 통계적 메타분석으로 최신 동향을 파악할 수도 있다.²⁾

EMC(Enterprise Content Management)는 디지털 정보자원을 체계적으로 수집·보존 및 큐레이션하여 무결성을 유지하고, 미래에도 접근 및 재사용할 수 있도록 디지털 큐레이션 및 보호에 기여하는 지역 도서관, 박물관 및 기록관 등에 다각적인 지원 사업을 시행하며, 세계정보문화유산의 디지털 보존에 대한 지원을 위하여 EMC 정보문화유산 기금프로젝트도 진행하고 있다.³⁾

디지털 큐레이션이 연구 및 재사용 가치가 있는 디지털 정보자원을 후손에게 이용하도록 하기 위한 방법으로 인식되고는 있지만, 과학자나 연구자들이 연구한 데이터들이 체계적으로 공유되거나 보존되지 않는다면 미래의 조사자, 연구자들에게 데이터가 재검토되고 재사용되는 것은 불가능하다. 이러한 문제점을 해결하기 위하여 미국과학재단(National Science Foundation)⁴⁾은 NSF 지원으로 수행된 연구로 생성·수집한 데이터를 최소한의 비용과 적절한 범위 내에서 다른 연구자에게 제공하도록 요구하고 있으며, 호주의 모나쉬 대학(Monash University)⁵⁾도 연구기간 동안의

1) National Academy of Sciences, *Free Executive Summary : Ensuring the Integrity, Accessibility, and Stewardship of Research Data in the Digital Age*, 2009.

〈http://www.nap.edu/html/12615/12615_EXS.pdf〉 [cited 2010. 2. 4].

2) Open Access news, 2009.

〈<http://earlham.edu/~peters/fos/2009/05/biological-data-repository-funded-by.html>〉 [cited. 2010. 2. 4].

3) 한국 EMC 홈페이지,

〈<http://korea.emc.com/leadership/digital-universe/information-heritage-project.htm>〉 [인용 2010. 2. 10].

4) NSF Home page, 〈<http://www.nsf.gov/>〉 [cited 2010. 2. 4].

5) Treloar Andrew, Groenewegen David and Harboe-Ree Cathrine., "The Data Curation Continuum : Managing Data Objects in Institutional Repositories," *D-Lib Magazine*, Vol.13, No.9/10(September/October 2007),

데이터 보존 및 큐레이션 문제를 다루기 위한 도메인, 데이터 저장, 큐레이션 경계의 개념을 제안하였다. 이는 데이터 객체를 평가, 기술, 제어하는 동시에 적합한 데이터 저장소로 이전할 수 있는 일련의 큐레이션 단계로서, 데이터는 개인적 연구자산에서 공유 연구자산 및 공공자산으로 이동하게 되는 것을 의미한다. 또한 영국에서도 연구를 통해 생성된 데이터에 영구적으로 접근하기 위해 디지털 큐레이션 센터(Digital Curation Centre)를 설립하였으며, 관련 도구와 정책을 수립하고 큐레이션에 관한 연구 및 교육 등을 활발하게 진행하고 있다.

디지털 정보자원을 중심으로 서비스하는 정보제공 기관에서는 연구기관, 연구자, 이용자에게 양질의 서비스를 제공하여야 하며, 디지털 데이터의 보존과 접근성 증대를 위한 지원도 아끼지 말아야 한다. 이를 위해서는 디지털 정보자원의 체계적인 관리 및 서비스를 통하여 장기보존에 필요한 큐레이션 환경 구축이 시급하다. 그러나 정책이나 절차 없이 기술적인 환경만으로 디지털 정보자원의 큐레이션을 이룰 수는 없으므로 디지털 정보자원의 장기보존과 큐레이션을 위해 필요한 조직, 대상 및 정책 등 기관의 지침이 우선적으로 갖춰져야 할 것이다.

본고에서는 디지털 큐레이션 정책을 수립하기 위하여 고려되어야 할 제반 프레임워크를 영국의 디지털 큐레이션 센터에서 시행하고 있는 큐레이션 정책과 실제로 데이터 공유 및 큐레이션에 대한 정책을 활용하고 있는 8개 연구재단의 사례를 분석하고, 디지털 큐레이션 생명주기를 기반으로 한 4가지 대범주와 14가지 소범주 정책 프레임워크를 개발하였다. 이는 향후 국내 각 기관이나 조직 차원에서 보존 및 큐레이션을 위한 정책과 전략을 수립하는 기초가 될 것으로 사료된다.

2. 연구 방법

디지털 정보자원의 장기보존 및 큐레이션을 시행하고 있는 네덜란드, 미국, 영국, 캐나다, 호주 등의 사례를 선행 연구자료 및 홈페이지를 통해 살펴보았다. 그중에서 가장 내실 있게 운영되고 있는 영국의 디지털 큐레이션 센터와 8개의 연구재단을 대상으로 큐레이션 정책과 지원서비스를 수집하여 디지털 큐레이션 생명주기 단계별 정의 및 큐레이션 정책을 분석하였다. 이를 기반으로 앞으로 국내에서 필요로 하는 디지털 큐레이션 생명주기 기반 정책 프레임워크를 개발하였다.

3. 선행 연구

국내의 경우, 이경남 외⁶⁾, 서혜란⁷⁾, 이수상⁸⁾ 등이 디지털 정보자원의 장기보존과 관리를 위한

⁶⁾ <<http://www.dlib.org/dlib/september07/treloar/09treloar.html>> [cited 2010. 2. 4].

⁶⁾ 이경남, 이소연, “전자기록의 장기적 보존과 관리를 위한 아젠다 개발,” 한국기록관리학회지, 제8권, 제1호(2008. 8), pp.211-234.

⁷⁾ 서혜란, “문화원형콘텐츠의 장기보존에 관한 연구,” 한국비블리아학회지, 제17권, 제2호(2006. 12), pp.65-82.

연구를 진행하였고, 이윤주 외⁹⁾는 전자기록의 보존과 정책에 대하여, 박현영 외¹⁰⁾는 아카이빙 정책에 대하여 보고한 바 있다. 그러나 디지털 정보자원의 보존 및 관리, 아카이빙 정책 및 전자기록의 보존과 정책은 비교적 활발하게 연구되어 왔지만 디지털 정보자원을 대상으로 장기보존과 접근 및 재사용을 위한 디지털 큐레이션에 관한 연구는 거의 전무한 상태다.

반면에 국외의 경우, 영국의 디지털 큐레이션 센터를 중심으로 프로젝트 및 저널 발행 등이 활발하게 이뤄지고 있으며, 큐레이션과 관련된 이벤트, 프로그램, 프로젝트와 관련 연구들은 홈페이지를 통해 제공되고 있다. Michael¹¹⁾은 영국의 주요 연구재단의 정책, 실무 및 견해에 대한 연구에서 큐레이션 대상으로 저널, 모노그래프, 컨퍼런스 자료, 논문, 데이터, 회색문헌 등의 연구 산출물이 원활하게 활용되도록 큐레이션 대상 자료들을 중심으로 정책 프레임워크를 개발하였다. Sarah¹²⁾는 영국의 주요 연구기관의 정책과 데이터관리에 관한 웹페이지 기반 조사를 통하여 큐레이션 정책에 대한 맵을 발표하였다. 또한 Raivo¹³⁾는 OECD 가입 국가의 연구데이터 공유를 위한 정책, 전략, 인프라 및 서비스를 각 나라별로 비교하고, 데이터 공유를 위한 정책 지원, 규정, 데이터 협력과 접근 서비스, 연구 진행 과정에 데이터 공유 지원서비스, 데이터 공유를 위한 연구 기술 등을 분석하였다.

II. 디지털 큐레이션 정책 및 사례

디지털 정보자원에 대한 보존과 관리, 공유 및 큐레이션을 위해 리포지터리와 데이터 센터, 데이터 관리와 공유, 기관 큐레이션 등의 정책을 검토하고, 영국의 디지털 큐레이션 센터와 8개의 연구기관을 중심으로 정책 사례를 고찰해 보았다.

- 8) 이수상, “디지털 아카이빙의 워크플로우와 보존처리 기술에 관한 연구,” 한국도서관·정보학회지, 제35권, 제3호(2004. 9), pp.119-138.
- 9) 이윤주, 이소연, “진본 전자기록의 장기보존을 위한 정책프레임워크 : InterPARES 성과물에 기초하여,” 기록학연구, 제19호(2009. 1), pp.193-249.
- 10) 박현영, 남태우, “디지털 아카이빙 정책에 관한 연구,” 제11회 한국정보관리학회 학술대회 논문집(2004. 8), pp.69-76.
- 11) Jubb Michael, “UK Research Funders’ Policies for the Management of Information Outputs,” *International Journal of Digital Curation*, Vol.2, No.1(January 2007), pp.29-48,
〈<http://www.ijdc.net/index.php/ijdc/article/viewFile/36/12>〉 [cited 2010. 2. 4].
- 12) Jones Sarah, *A report on the range of policies required for and related to digital curation*, DCC Policies report, Version 1.2, 2009, pp.1-34,
〈http://www.dcc.ac.uk/sites/default/files/documents/reports/DCC_Curation_Policies_Report.pdf〉 [cited 2010. 3. 16].
- 13) Ruusalepp Raivo, *A comparative study of international approaches to enabling the sharing of research data*, Version 1.6, 2008,
〈http://www.jisc.ac.uk/media/documents/programmes/preservation/national_data_sharing_report_final.pdf〉 [cited 2010. 2. 4].

1. 리포지터리와 데이터 센터 정책

리포지터리와 데이터 센터에서는 큐레이션을 지원하기 위해 적절한 분산형 기본시설을 구축하는 것이 중요하다. 리포지터리의 구성요소로는 실험실 리포지터리, 기관 리포지터리, 주제 또는 학문분야 리포지터리, 데이터베이스와 데이터 센터 등이 포함된다.

큐레이션 정책을 수립하려면 큐레이션 대상 자료를 어떻게 저장하고 접근할지에 대한 기본적인 사항 등을 비롯하여 장기 보존할 수 있도록 과정과 절차에 대한 세부 설명이 명시되어야 한다.

OpenDOAR(The Directory of Open Access Repositories)¹⁴⁾는 각각의 리포지터리 메타데이터, 재사용, 콘텐츠, 제출 및 보존 정책을 제공하며, 리포지터리 관리자는 정책을 선택 입안하고 동시에 다양한 온라인정책 툴도 서비스한다.

UKDA(The UK Data Archive)¹⁵⁾ 데이터 센터는 큐레이션 역할과 책임을 입수 전, 입수, 아카이빙, 데이터 관리, 접근, 관리 기능 등 세부항목으로 정의하여 큐레이션 생명주기를 포함한 보존 계획과 전략, IT 아키텍처, 보안, 협력, 기금 및 자원 계획 등으로 구분하였다.

SHERPA JULIET 서비스¹⁶⁾는 출판물에 대한 정책을 통하여 오픈엑세스 출판과 아카이빙에 대한 세부 정보를 제공하며, RoMEO 서비스¹⁷⁾는 출판사의 저작권과 셀프 아카이빙 정책에 필요한 데이터베이스를 Green, Blue, Yellow, White 단계로 구분하여 Post-Print 또는 Pre-Print 여부를 제공한다.

이상의 기관들은 정책 개발을 위한 리포지터리 지원으로 OpenDOAR와 RSP(Repositories Support Project)¹⁸⁾, SUETr¹⁹⁾ 교육프로그램 및 리포지터리 개선 계획을 활용하여 연구 활동을 지원한다.

2. 데이터 관리와 공유 정책

연구결과물의 주체가 되는 기관과 리포지터리 보존 정책과 관련된 데이터 관리 계획이 필요하고,

14) OpenDOAR은 영국의 기관 리포지터리 추진 프로그램인 SHERPA(Securing a Hybrid Environment for Research, Preservation an Access)에서 운영하는 기관 리포지터리 디렉토리로 "OpenDOAR"에 등록된 Repository Entry가 1,000건을 돌파(2007. 7. 21)하고 기관 리포지터리의 확산은 세계적인 추세이다. OpenDOAR Home page, <<http://www.opendoar.org/tools/en/policies.php>> [cited 2010. 2. 4].

15) Woollard Matthew, *UK Data Archive Preservation Policy*, Version 03.10, 2009, <<http://www.data-archive.ac.uk/news/publications/preservationpolicy.pdf>> [cited 2010. 2. 4].

16) JULIET and Support for Open Access to Research Home page, <<http://www.sherpa.ac.uk/juliet/>> [cited 2010. 2. 4].

17) Romeo Home page, <<http://www.sherpa.ac.uk/projects/sherparomeo.html>> [cited 2010. 2. 4].

18) Repositories Support Project Home page, <<http://rsp.ac.uk/repos/rules>> [cited 2010. 2. 4].

19) SUETr Home page, <<http://www.suetr-rp.wikispaces.com/>> [cited 2010. 2. 4].

이러한 관리 계획은 큐레이션 인프라를 제공하는 순간부터 발생하게 된다. 연구 데이터의 큐레이션 단계에서 연구결과물을 제출하는 연구자들은 지적재산권에 따른 윤리적 고려사항 및 큐레이션에 사용되는 방법론, 표준, 데이터의 유지, 접근 및 재사용 방법 등과 같은 세부사항이 큐레이션을 제안하는 단계에서부터 포함되기를 기대하고 있다.

NERC와 ESRC의 경우는 데이터 센터 직원은 연구자들과 연계하여 제안 단계에서부터 참여하여 연구결과물을 생성하고 관리한다. 또한 UKDA 센터에서는 데이터의 관리와 공유²⁰⁾를 위하여 UKDA의 RELU(Rural and Economic Land Usage) 프로그램²¹⁾으로 맞춤형 데이터 관리 지원 서비스를 제공한다.

3. 기관 큐레이션 정책

기관에서의 레코드 관리에 대한 디지털 큐레이션 정책은 기관마다 기술적, 행정적 수준에 적합한 작업 프레임, 데이터 정책부서, 데이터 관리 계획에 대한 지침이 요구되며, 각 기관 및 부서 간 상호 협력을 통하여 발전할 수 있다.

2002년에 영국의 7개 연구위원회가 공동으로 설립한 RUCK(Research Councils UK)²²⁾는 기관 차원에서 큐레이션 정책이 운영되어야 하고, 기관이 연구기금으로부터 지원 받은 데이터를 보유, 소유 및 접근할 수 있도록 요구하고 있다. 특히, 다양한 과학 분야에 대한 연구와 교육을 지원하고, 영국의 산업발전을 위해 연구자와 연구 기반 시설을 보유하고 기술 이전이 차질 없이 진행되도록 임무를 수행하고 있다. 또한 신규 연구과제도 발굴하거나 기획하는 업무도 한다.

Beagrie 정책²³⁾은 기관의 보존 정책을 생성, 계획하기 위한 프레임워크를 제공하여 보존 정책과 정보, 도서관과 레코드 관리 전략 사이의 협력 관계에 대하여 제시하고 있다.

〈그림 1〉은 디지털 큐레이션 생명주기 단계별 정책과 이해관계자 사이에서 리포지터리와 데이터 센터, 데이터 관리와 공유, 기관 큐레이션 정책 영역별 관계를 표시한 매핑이다.

리포지터리와 데이터 센터의 주요 정책 영역은 평가, 수집, 메타데이터, 저장, 보존/변환, 접근, 사용 및 재사용 영역이고, 기관 큐레이션의 주요 정책 영역은 데이터 생성, 메타데이터, 평가 및 선정, 저장, 보존, 접근·사용 및 재사용 영역이며, 데이터 관리와 공유의 주요 정책 영역은 개념화, 데이터 생성, 메타데이터, 저장, 보존, 접근 및 재사용에 있다.

20) UKDA Home page, <<http://www.data-archive.ac.uk/sharing/sharing.asp>> [cited 2010. 2. 4].

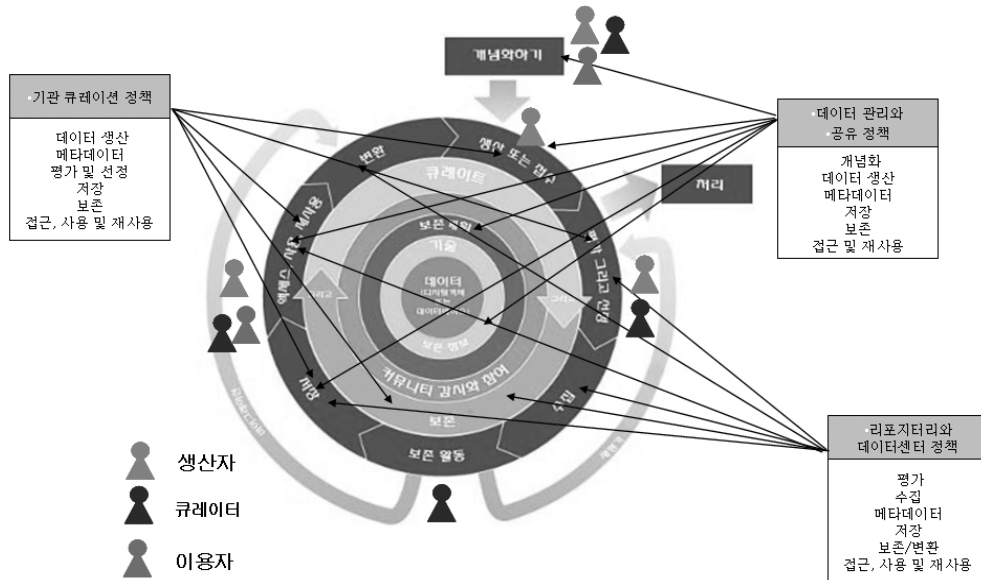
21) UKDA relu Home page, <<http://www.data-archive.ac.uk/relu>> [cited 2010. 2. 4].

22) RUCK Home page,

<<http://www.rcuk.ac.uk/cmsweb/downloads/rcuk/reviews/grc/goodresearchconductcode.pdf>> [cited 2010. 2. 4].

23) Neil Beagrie et al., *Digital preservation policies study*, Charles Beagrie Limited, 2008,

<http://www.jisc.ac.uk/media/documents/programmes/preservation/jiscpolicy_plfinalreport.pdf> [cited 2010. 2. 4].



〈그림 1〉 큐레이션 정책과 이해관계자 간의 관계 매핑²⁴⁾

이해관계자인 생산자, 큐레이터, 이용자의 역할과 큐레이션 정책이 상당부분 중복되는데, 생산자는 데이터 센터에서 장기 큐레이션의 책임이 따르고, 개념화, 평가, 선택은 다양한 이해관계자의 투입(input)이 요구된다. 생산자에게는 큐레이터 및 이용자 사이에 커뮤니티가 형성되어야 하고, 큐레이터와 이용자 사이에도 커뮤니티가 요구된다. 또한 데이터의 위험을 최소화하도록 큐레이션 생명주기를 통하여 이해관계자 모두는 디지털 정보자원의 접근, 이용에 협력해야 한다.

4. 디지털 큐레이션 정책 사례

가. DCC(Digital Curation Centre)의 큐레이션 정책

디지털 정보자원은 생성부터 보존 및 변환에 이르기까지 생명주기 관점에서 단계별 세부계획이 명시되어야 한다. 디지털 자원은 생명주기에 따른 훼손, 파손 및 노후화 등 기술의 변화에 민감하고, 생명주기에 영향을 미치는 각 단계별 활동은 연속적으로 디지털 자원의 보존 및 관리 등에 영향을 미친다.

JISC(Joint Information Systems Committee)²⁵⁾는 기본적인 보존정책 가이드를 만들었고, JISC

24) Higgins Sarah, *Applying the DCC Curation Lifecycle Model*, 2009,

〈http://www.fsd.uta.fi/iassist2009/presentations/B3_Higgins.ppt〉 [cited 2010. 2. 4].

25) 영국 합동정보시스템위원회(Joint Information Systems Committee, JISC) : 영국의 고등교육 즉, 대학의 강의, 학습, 연구, 행정 등의 영역에서 정보와 통신기술의 효율적인 사용을 위한 각종 프로젝트를 수행하는 조직. JISC

의 요구로 디지털 큐레이션 정책에 대한 연구가 시작되었다. DPC(The Digital Preservation Coalition)와 DCC에서는 2006년 7월에 '디지털 큐레이션과 보존에 관한 정책' 포럼²⁶⁾에서 디지털 정보자원의 효과적인 관리, 발견, 재사용과 관련된 정책을 논의한 결과, DCC의 큐레이션 정책²⁷⁾은 글래스고 대학에서 정책 연구를 수행하며, 큐레이션을 지원하기 위한 권장사항 제공 및 지속적인 관련 연구를 진행하게 되었다.

DCC는 글래스고 대학과 협력하여 글래스고 대학의 HATII 프로젝트에 의해 DAF(Data Audit Framework)²⁸⁾를 개발하고 구현하였다. DAF에는 방대한 양의 데이터가 생성되어 있으므로 연구결과물에 대한 공식적인 큐레이션 전략을 갖고 있지 않은 문제점을 해결하는 동시에 기관 연구자들의 데이터 관리 및 유지에 대한 정책을 수립하기 위해 만들었다. DAF는 감사 방법과 온라인 툴, 레지스트리를 생성하고, 큐레이션 관련 서비스를 지원하고 있다.

나. 영국의 8개 연구재단의 정책 사례

(1) AHRC²⁹⁾ - 예술 및 인문학 연구 위원회

AHRC(Arts and Humanities Research Council)의 정책은 ① 데이터 공개와 이용을 위해 시간 제한을 두었다. 전자자원 또는 데이터 세트는 프로젝트 완성 후 3개월 이내에 ADS(Archaeology Data Service)³⁰⁾에 제공하고, 저장 후 최초 프로젝트 이후 3년 동안 접근할 수 있도록 유지해야 한다. ② 신청자가 주요 전자자원을 생성하기 위한 계획을 만들려면 제안 단계에서 프로젝트 관리, 데이터 개발, 인프라 지원, 데이터 보존 및 유지, 접근, 저작권과 관련된 이슈 등 기술적인 측면을 포함하여야 한다. ③ 모든 연구결과가 접속 가능한 리포지터리에서 이용되어야 하고, 전자자원은 최소 3년 동안 유용하도록 유지되어야 하며, 데이터 생성과 관리를 위해 기술적 사항이 필요하다. ④ 큐레이션 정책에 관한 준수사항은 모니터 되지 않으며, 이러한 요구를 수행하지 않는 것에 대한 제재도 없다.

(2) BBSRC³¹⁾ - 생명과학 및 생물과학 연구 위원회

BBSRC(Biotechnology and Biological Sciences Research Council) 정책은 ① 연구결과물의

Home page, <<http://www.jisc.ac.uk>> [cited 2010. 2. 4].

26) DPC Home page,

<<http://www.dpconline.org/events/policies-for-digital-curation-and-preservation.html>> [cited 2010. 2. 4].

27) DCC Home page, <<http://www.dcc.ac.uk/resources/policy-and-legal>> [cited 2010. 2. 4].

28) Jones Sarah, "The Data Audit Framework : A First Step in the Data Management Challenge," *The International Journal of Digital Curation*, Vol.3, No.2(November 2008), pp.112-120, <<http://www.ijdc.net/index.php/ijdc/article/viewFile/91/62>> [cited 2010. 2. 4].

29) AHRC Home page, <<http://www.ahrc.ac.uk/Pages/default.aspx>> [cited 2010. 2. 4].

30) ADS Home page, <<http://ads.ahds.ac.uk/>> [cited 2010. 2. 4].

31) BBSRC Home page, <<http://www.bbsrc.ac.uk/>> [cited 2010. 2. 4].

공유를 위해 시간제한을 두고, 출판물은 우선적으로 저장해야 하며, 데이터는 즉시적으로 이용 가능하고, 프로젝트 완성 후 데이터는 10년 동안 유지되어야 한다. ② 연구자들은 제안서에 데이터 영역, 유형 및 포맷, 표준 및 메타데이터, 부가사용, 데이터 공유 방법 및 공개시간 등의 세부사항을 포함한 데이터 공유 계획을 제출해야한다. ③ 위원회에서는 데이터 공유 증진 및 공개적으로 자료에 접속토록 하기 위한 활동을 하고 데이터 공유를 위한 재정적 지원과 출판 리포지터리를 계속적으로 유지할 수 있도록 한다. ④ 제안된 데이터 관리와 공유 계획에 대한 사항은 재정적 보고서의 평가 과정을 통해 모니터링 되고, 향후 제안을 평가할 때 참고한다.

(3) EPSRC³²⁾ - 공학 및 자연과학 연구 위원회

EPSRC(Engineering and Physical Sciences Research Council) 정책은 ① 시간제한 규정만 있고, 데이터관리 및 공유에 대한 원칙을 제공하도록 요구하고 있지 않다. ② 연구결과물에 대한 접근에 대해서는 RUCK 성명서를 따른다. ③ 주요 데이터의 보안과 저장은 소유권이 있는 기관의 통제 하에 저장되어야 하고, 모니터링은 하지 않는다.

(4) ESRC³³⁾ - 경제 및 사회 연구 위원회

ESRC(Economic and Social Research Council) 정책은 ① 시간제한은 데이터와 출판물 공개를 위한 조항이며, 데이터가 소유권이 만료되는 최소 3개월 이내에 ESDS(Economic and Social Data Service)에 제공되어야 한다. ② 연구제안 단계에서 어떤 연구결과물이 나올지, 장기간 이용될 수 있을지에 대하여 고려하게 하고, 기존의 데이터 조사를 포함한 아카이빙 및 잠재 이용자들을 위해 조사한다. ③ 연구결과물에 빨리 접근할 수 있도록 출판물 리포지터리와 데이터 서비스를 제공한다. ④ 연구자들은 연구결과물이 생성되고 적절하게 유지되도록 지침이 필요하므로 위원회는 장기 큐레이션에 관한 규정을 리포지터리와 데이터 아카이브 서비스를 통해 보여준다. ⑤ 정책에 따른 모니터링을 한다.

(5) MRC³⁴⁾ - 의학 연구 위원회

MRC(Medical Research Council) 정책은 ① 출판물이 공식적인 데이터와 최종 출판 후 6개월 이내에 접근할 수 있기를 기대한다. 또한 연구가 완성된 후 10년간 안전하게 유지되어야 하고, 주요 연구를 위한 데이터 사용 기간이 정의되어야 한다. ② 연구자에게는 제안 단계에서부터 데이터 공유와 보존 전략을 통하여 향후 그들의 연구 자원에 대해 고려하도록 한다. ③ 출판물에 대한 접

32) EPSRC Home page, <<http://www.epsrc.ac.uk/default.htm>> [cited 2010. 2. 4].

33) ESRC Home page, <<http://www.esrc.ac.uk/ESRCInfoCentre/index.aspx>> [cited 2010. 2. 4].

34) MRC Home page, <<http://www.mrc.ac.uk/index.htm>> [cited 2010. 2. 4].

근을 제한하지 않고 데이터에 접속할 수 있도록 한다. ④ 장기 큐레이션 및 연구 결과에 따른 보존에 대한 권고사항을 제공하고 있으며, 이러한 정책에 관한 모니터링은 아직 수행하고 있지 않다.

(6) NERC³⁵⁾ - 자연 환경 연구 위원회

NERC(Natural Environment Research Council) 정책은 ① 연구결과물은 저장 및 제공되어야 하며, ② 데이터 생성과 유지에 관한 계획과 서류 작성 및 지적 소유권 수립에 대한 절차를 진행한다. 또한 데이터 생성, 저장, 백업, 아카이브, 접근 및 재사용에 관한 역할과 책임에 대해 명시한다. ③ 출판물과 데이터는 리포지터리와 데이터 센터를 통해 접속 가능해야하고, 데이터가 재사용될 수 있도록 최소한의 질적 기준이 마련되어야 한다. ④ 장기 큐레이션에 대한 인프라를 제공하며, ⑤ 출판 정책의 준수는 향후 재사용을 위해 고려한다.

(7) STFC³⁶⁾ - 과학 및 기술 시설 위원회

STFC(Science and Technology Facilities Council) 정책은 ① 데이터를 위한 시간제한은 없다. ② 데이터 관리나 공유에 관한 규정이 마련되어 있지 않고 장기간 큐레이션을 위한 정책 또한 갖고 있지 않다. ③ 공개접속 정책을 준수하거나 모니터링을 하고 있지 않다.

STFC는 다른 기관에 비하여 전체적으로 데이터 접속, 공유 및 큐레이션을 위한 정책에 대한 규정이나 지침 등이 미비한 편이다.

(8) The Wellcome Trust³⁷⁾

Wellcome Trust 정책은 ① 출판물과 데이터 양쪽 모두에 시간제한에 대한 조건을 제시하며, 연구된 결과물은 최종 출판 후 6개월 이내에 저장되어야 한다. 또한 10년간 데이터를 안전하게 유지해야 한다. ② 데이터 관리와 공유 계획을 제출하고, 지침서는 데이터의 질과 표준, 공동 데이터 리포지터리 사용, 지적 재산권, 연구 참여자의 보호, 장기 보존과 지속성에 대하여 명시하고 있다. ③ 연구결과물의 오픈엑세스를 증진시키는 역할을 하며, 어떤 제한 없이 데이터를 이용할 수 있도록 하고 있다. ④ 연구자원의 생성과 유지를 위해 데이터의 저장과 처리를 위한 책임과 절차에 대한 지침을 제공하고 있다. ⑤ 연구결과물에 대한 오픈엑세스에 대한 정책의 준수는 모니터 되며, 데이터 공유의 법적 소유권에 대한 사항은 모니터링 하지 않는다.

35) NERC Home page, <<http://www.nerc.ac.uk/>> [cited 2010. 2. 4].

36) STFC Home page, <<http://www.scitech.ac.uk/>> [cited 2010. 2. 4].

37) Wellcome Trust, *Policy on data management and sharing*, <<http://www.wellcome.ac.uk/About-us/Policy/Policy-and-position-statements/WTX035043.htm>> [cited 2010. 2. 4].

이상과 같이, 영국의 8개 연구재단에 의한 큐레이션과 서비스 지원은 정책 범위, 큐레이션 규정, 지원 규정으로 구분되며, 지원서비스는 출판물, 데이터, 시간제한, 데이터 계획, 접근/공유, 큐레이션, 모니터링, 지침, 리포지터리, 데이터센터 항목으로 세분화된다(표 1 참조).

〈표 1〉 영국의 8개 연구재단의 큐레이션 정책과 지원 서비스³⁸⁾

구분 연구재단	정책 범위		큐레이션 규정					지원 규정		
	출판물	데이터	시간 제한	데이터 계획	접근/공유	큐레이션	모니터링	지침	리포지터리	데이터 센터
AHRC - 예술 및 인문학 연구 위원회	●	●	●	●	●	●	○	●	○	●
BBSRC - 생명과학 및 생물과학 연구 위원회	●	●	●	●	●	●	●	●	●	●
EPSRC - 공학 및 자연 과학 연구 위원회	●	○	●	○	●	○	○	○	○	●
ESRC - 경제 및 사회 연구 위원회	●	●	●	●	●	●	●	●	●	●
MRC - 의학 연구 위원회	●	●	●	●	●	●	○	●	●	●
NERC - 자연 환경 연구 위원회	●	●	●	●	●	●	●	●	●	●
STFC - 과학 및 기술 시설 위원회	●	○	●	○	●	○	○	○	●	●
Wellcome Trust	●	●	●	●	●	●	●	○	●	●

표시 : ● 전체포함 ● 부분포함 ○ 미포함

각 항목에 따른 연구재단의 큐레이션 정책과 서비스에 대하여 분석해 보았다.

□ 출판물

출판물에 대한 연구재단의 정책은 2005년 7월 RUCK³⁹⁾의 규정으로 출판비용을 어떻게 정해야 하는지에 대한 고려사항이 조정되었으며, NERC의 경우는 부가조건을 포함해서 정책을 세우고 있다. EPSRC는 부분적으로 출판물에 대한 정책을 갖고 있다.

□ 데이터

전자자원 관리나 자원 접근에 관한 데이터 정책이나 규정은 다각적으로 이루어지고 있으며, 적용범위도 다양하다. 그러나 EPSRC와 STFC는 데이터 정책을 갖고 있지 않다.

38) Sarah Jones, *op. cit.*, p.8.

39) Brighton Robin, *Open access to research output final report to RUCK*,
 〈<http://www.rcuk.ac.uk/cmsweb/downloads/rcuk/news/oareport.pdf>〉 [cited 2010. 2. 4].

□ 시간제한

콘텐츠의 접근 및 검색결과에 대한 접근으로 출판 결과 최소 6개월 이내에 접근할 수 있도록 요구하고 있으나, ESRC는 3개월 이내에 데이터가 저장되고, 프로젝트를 착수함과 동시에 데이터 센터와 관계를 유지한다. BBSRC, MRC, Wellcome Trust는 데이터가 연구 프로젝트 완성 후 10년간 안전하게 보호하는 규정을 두었으나 STFC와 EPSRC는 시간제한에 대한 규정을 갖고 있지 않다.

□ 데이터 계획

데이터를 만들고, 관리하고, 적용을 위한 공유 등을 제안 단계에서 연구자에게 접근이나 운용 및 장기간 큐레이션 하도록 요구한다. AHRC, ESRC, NERC는 어떻게 자원이 만들어질 것인지, 장기간 보존될 수 있을 것인지에 대한 규정을 요구하고, BBSRC, MRC, Wellcome Trust는 잠재적 연구자원의 데이터 공유에 비중을 두는 반면에 EPSRC와 STFC는 데이터 공유를 달성하기 위한 계획을 요구하지 않는다.

□ 접근/데이터 공유

RUCK 규정에 서명한 후원자들은 연구결과에 언제라도 접근할 수 있고, 연구결과물도 올릴 수 있다. MRC, Wellcome Trust는 자유롭게 복사할 수 있고, 텍스트나 데이터 개발과 같은 목적으로 재사용할 경우에 접근이 가능하도록 소스 데이터와 출판물을 연결하였다. BBSRC, MRC, Wellcome Trust는 데이터 공유에 자유롭고, ESRC, NERC는 데이터 센터를 지원하여 데이터 공유를 활성화하고 있다. AHRC도 접근을 허용하고 교차검색을 지원한다. STFC는 데이터 공유에 관해 분명한 규정을 제공하고 있지 않고, EPSRC 관계자들은 데이터 공유의 필요성에 대해 관심을 갖고 있지 않다.

□ 큐레이션

AHRC, BBSRC, EPSRC, MRC, NERC, Wellcome Trust는 큐레이션 주기를 다양한 측면에서 고려하고 있으나, EPSRC는 데이터가 10년간 저장되어야 한다는 조건을 설정하였다. STFC는 큐레이션을 설명하는 공식적인 요구가 없다.

□ 모니터링

NERC, Wellcome Trust는 오픈액세스 정책을 모니터링을 하고, BBSRC는 데이터 관리와 데이터 공유 계획을 고수할 수 있도록 모니터링 하여 미래의 제안을 위한 고려사항으로 참고한다. 그러나 ESRC는 다른 연구재단의 규정에 대해 모니터를 하지 않고 큐레이션 정책에 따르지 않았

을 때의 불이익에 관해서도 명시하고 있지 않다.

□ 지침

서비스를 지원하고 안내하기 위한 지침으로 연구자들을 위한 온라인 정책을 제공한다. MRC는 실제적인 안내와 데이터 공유 세트를 제공하고 BBSRC는 온라인상에 안내를 제공하고 있지 않지만 관계된 기준과 정보에 대해 설명하며 주요 사이트도 제공한다. EPSR, STFC, Wellcome Trust는 지침에 대한 정책이나 웹사이트를 제공하고 있지 않다.

□ 리포지터리

AHRC, EPSRC는 연구결과물에 대한 리포지터리를 제공하지 않지만, 대부분의 연구재단은 그들이 후원하는 연구자들을 위해 리포지터리 제공한다.

□ 데이터센터

미출판 된 전자자원 또는 데이터 생성에 대한 데이터 센터의 조항에서 데이터 센터의 충분한 데이터 서비스를 갖고 있는 기관은 거의 없다. 다만 ESRC, NERC는 데이터 서비스를 제공하고 지원하고 있다.

이상과 같이 각 항목에 따른 큐레이션 정책과 서비스에서 ESRC, NERC의 경우에 출판물과 데이터에 정책을 모두 포함시킬 뿐 아니라 큐레이션 정책과 규정 지원에 대한 항목들도 모두 포함하고 있음을 알 수 있다. 반면에 EPSRC, STFC의 경우에는 큐레이션에 대한 정책 및 규정을 반영하고 있지 않음을 알 수 있다.

Ⅲ. 디지털 큐레이션 정책 프레임워크 개발

1. 디지털 큐레이션 정책 프레임워크 구조

큐레이션 정책에는 데이터의 획득, 관리, 보존, 아카이빙을 위한 국가적이고 국제적인 표준을 따르는 큐레이션 실무 절차를 거쳐, 발견과 접근의 통제, 데이터의 공유와 재사용을 촉진하는 과정, 도구 등이 필요하다. 데이터 센터와 리포지터리 시설을 기본으로 출처의 표현과 진본성에 대한 보증을 위한 방법론 등도 고려되어야 한다.

또한 큐레이션은 현재 진행 중인 연구에 관한 동료 연구자간의 비형식적인 교류와 교육, 형식적

규정과 윤리적 과정을 준수하는 등의 실제 사례를 기반으로 이루어지기 때문에 구체적인 큐레이션 실무에서는 데이터의 구조와 규모, 소유권뿐만 아니라 문화의 다양성과 연구 방법론도 고려해야 한다.

본고에서는 DCC의 큐레이션 정책과 영국의 8개 연구재단의 지원 서비스 등의 요소를 기반으로 큐레이션 생명주기 단계에 따른 정책 프레임워크를 개발하였다. 디지털 큐레이션 정책 프레임워크 구조는 디지털 큐레이션을 위한 원칙, 정책 수립, 생명주기 기반 큐레이션 정책, 지원 규정의 4가지 범주 아래에 14개의 하위범주를 두었다(그림 2 참조).



<그림 2> 디지털 큐레이션 정책 프레임워크 구조

디지털 큐레이션을 위한 원칙 및 지침 범주에는 목적, 역할 및 책임의 하위범주를 설정하고, 정책 수립 범주에는 큐레이션 범위, 법적 권리 및 윤리적 이슈, 권장사항을 하위범주로 구성하였으며, 생명주기 기반 큐레이션 정책의 범주에는 관리 계획, 수집 및 개발, 평가 및 선정, 저장 및 관리, 접근, 공유 및 재사용, 변환을 하위범주로 하였고, 지원 규정 범주에는 교육, 홍보 및 가이드, 모니터링을 하위범주로 설정하였다.

2. 생명주기 기반 큐레이션 정책 프레임워크 개발

미국 의회도서관에서 소장 자료의 디지털화를 위해 기본 목표를 정비할 당시에 디지털콘텐츠의 생명주기를 고려하여 운영정책을 확립⁴⁰⁾ 하였듯이, 각 기관마다 기관의 역할과 사명에 따라 특성은 다르겠지만 DCC의 생명주기를 기반으로 국내 실정에 맞게 기본적인 큐레이션 정책 프레임워크를

40) Library of Congress, *Library of Congress Strategic Plan*(Washington DC : Library of Congress, 2006).

개발하고자 한다.

큐레이션 정책은 기존의 업무와 조화를 이루어 시행되어야 하는데, 기관 자체뿐만 아니라 기관 간의 새로운 파트너십과 협력관계를 구축하고, 기관 이용자의 요구를 파악하여 정책이 개발되어야 한다. 큐레이션 활동 관련 이해관계자들과 협력하고, 큐레이션 전략 마케팅을 개발하여 정책 범위, 큐레이션 규정 및 지원 정책 등도 명시되어야 한다.

또한 기관별 디지털 정보자원의 큐레이션 정책 범위에 따른 수집 및 관리 정책이 수립되어야 하며, 디지털 정보자원의 공개 범위, 시기, 이용 제한 등에 따른 정책 및 규정도 명시되어야 한다. 뿐만 아니라 저작권 관련 지침 및 법적, 제도적 지원 장치가 마련되어야 하고, 연구자가 참여할 수 있도록 연구 성과 정보 등의 제출 방식도 간소화될 수 있도록 시스템을 개발해야 한다. 나아가 다른 시스템과 상호 운용성을 고려하여 국제표준을 준수하고, 디지털 정보자원의 장기 보존을 위한 방안 또한 마련해야 한다.

디지털 정보자원의 효과적인 관리를 위해서는 다음과 같은 원칙이 필요하다.

- 첫째, 연구자, 연구기관 및 관련 부서의 역할과 책임이 분명하게 정의되어야 한다. 연구 데이터의 이용자와 생성자의 책임을 원칙에 따라 인식하고 수행하도록 규정을 체계적으로 마련한다.
- 둘째, 디지털 정보자원은 세계적인 표준에 맞게 만들어져야하고, 수집되어야 하며, 이러한 선택 과정은 다른 사람들이 이용할 수 있도록 질적인 보장이 포함되어야 한다.
- 셋째, 디지털 정보자원은 검색이 용이해야 하며, 최대한 접속에 필요한 이용 가능한 환경을 제공하고, 데이터 생성자나 수집가를 위한 권리를 보호하는 법적·윤리적 제도를 마련해야 한다.
- 넷째, 디지털 정보자원에 접속을 제공하고 유지하기 위한 모델과 메커니즘은 공개적으로 사용되며, 비용 면에서 효율적이어야 한다.
- 다섯째, 장기적으로 보존 및 재사용을 위한 디지털 정보자원이 보호되어야하고, 현재 및 미래 세대를 위해 접근이 가능해야 한다.

이상의 효율적인 관리 원칙을 토대로 큐레이션 정책 개발 방법을 다음과 같이 구성하여야 한다.

- 첫째, 큐레이션을 위한 기관의 사명 및 목적, 범위를 정확히 명시한다.
- 둘째, 작업을 수행할 책임자에 대한 역할 및 책임에 대한 정의가 필요하다.
- 셋째, 상호운용을 위한 표준을 마련한다.
- 넷째, 보존 및 접근 재사용에 관한 기준이 필요하다.
- 다섯째, 이익 창출과 비용 효과에 대한 기준을 세워야한다.
- 여섯째, 각 기관에서 보유하고 있는 컬렉션을 파악하고, 특성을 확인하여 우선순위를 결정한다.

이상과 같은 효율적인 관리 원칙과 큐레이션 정책 개발 방법은 기존의 현장에 변화를 수반하게

되는데, 기존의 워크플로우 구조를 기반으로 정책을 마련해야하며 부수적으로 발생하는 인프라와 서비스도 지원해야 한다. 모든 정책에는 항목에 따른 목적이 명확하게 명시하여야 하고, 이를 시행하기 위하여 담당자의 책임과 임무에 대해서도 기술되어야만 비로소 성립하게 된다.

이를 시행하기 위하여, 국내에서 적용이 가능한 세부적인 큐레이션 정책 수립에 필요한 디지털 큐레이션을 위한 원칙 및 지침, 정책 수립, 생명주기 기반 큐레이션 정책, 지원 규정을 항목별로 제시하면 아래와 같다.

가. 디지털 큐레이션을 위한 원칙 및 지침

(1) 목적

데이터의 큐레이션과 보존을 위한 경영 계획과 관리 계획에 대한 기관의 사명 및 기관 전체, 한 부서 또는 연구 프로젝트에 따른 정책 및 연구결과물의 유형에 대한 정의, 데이터의 최소 기준, 안정적인 품질 평가, 비용 절감 효과 등을 명시한다.

(2) 역할 및 책임

작업을 수행할 책임자 및 생명주기 각 단계별 역할 및 정의가 있어야한다. 현재 진행되고 있는 활동과 정책을 연결하고, 기존의 워크플로우와 직원의 능력을 조정한다. 데이터의 이용자와 고객, 재사용자 및 재정지원 기구 모두는 큐레이션에 대한 책임을 수반한다.

나. 정책 수립

(1) 큐레이션 범위

적절한 서비스와 인프라를 동반하고, 기존 정책의 보완 방법과 광범위한 맥락에서 고정시키는 방법이 필요하다. 범위 활동으로 정책에 대한 요구를 파악하고, 제안된 변경사항에 대해 알려야 하며, 정책을 정기적으로 검토하고 적절하게 유지되도록 지속적으로 살펴보아야 한다.

(2) 법적 권리 및 윤리적 이슈

디지털 자료 보존에 대한 책임과 범위를 확정하기 위하여 디지털 자료 보존에 관한 표준화를 증진시키고 디지털 자료의 위험 관리를 위하여 보존 정책, 법률적인 사항 등을 명시하여 데이터 공유 이전에 법적인 절차가 완료되어야 한다. 디지털 데이터의 장기보존에 대한 법률적 제약은 저작권 문제가 비중을 크게 차지하는데, 데이터의 소유권자가 재사용을 허용하지 않는 경우에 저작권 위반이 될 수 있고, 또한 보존을 목적으로 데이터를 복제하는 행위도 저작권자의 승인이 없다면 저작권법에 위배된다. 그러므로 저작권과 라이선스, 법적·윤리적으로 데이터를 보호하기 위한 법률적인 문제를 해결하는 것이 우선되어야 한다.

(3) 권장 사항

새로운 정책에 대한 호응을 얻고, 연구자의 인식, 시스템에 대한 보상 및 테스트 단계 등을 실행해야 할 큐레이션 정책을 개발하고 계획하기 위해 관련 기관에서 제공하고 있는 정책 도구와 자원 및 지침 등을 참고하고, 정책은 맥락에 유념하여 데이터 관리 계획에 대한 보완과 큐레이션을 위해 관련 리포지터리 기관과 협력하여 작업이 이뤄져야 한다. 또한 기존의 워크플로우 구조에 큐레이션을 삽입하여 사용할 수 있도록 기존의 주제사서가 큐레이터로서의 역할을 지원할 수 있게 한다.

다. 생명주기 기반 큐레이션 정책

생명주기 단계에 따른 큐레이션 정책에 포함될 내용은 각 기관의 현장 상황에 따라 적용하는 것이 바람직하다.

(1) 데이터 큐레이션과 보존을 위한 관리 계획

보존 정책의 실행을 위한 전략 및 실행을 지원하는 도구로서의 표준은 정책과 일관성 있게 수립해야 한다. 디지털 보존 목표는 현재와 미래를 위해 보유한 데이터를 일정기간 동안 이용 가능해야 하고, 진본성이 보장되고 신뢰할 수 있어야 한다. 또한 미래 잠재 이용자 커뮤니티에 적합한 방식으로 데이터를 배치해야 한다.

보존 계획으로 어떤 데이터가 언제 누구에 의해 생성되어 구조화되었는지 여부와 데이터가 어디에서 유지되고 연구 프로젝트의 마지막 단계에서 누가 데이터 유지를 위한 재정지원을 할 것인지 등이 명시되어야 한다.

보존 활동으로는 디지털 데이터를 식별, 기술, 조직, 보존, 등록하고, 용어, 개념, 권한 관리 및 접근 제어 등을 성문화할 수 있는 소프트웨어도 명시하여야 한다.

(2) 데이터 수집 및 개발

기관의 사명에 맞는 수집 및 개발 정책이 필요하므로 디지털 자료의 캡처부터 아카이빙 포맷을 위한 표준 및 정보 출처, 맥락에 관한 정보 기준이 마련되어야 한다. 또한 데이터의 진본성, 정확성, 표현 가능성, 지속성 등이 확실하도록 데이터 구조와 데이터 품질, 인용, 어노테이션, 출처 정보에 대한 표준도 명시한다.

(3) 평가 및 선정

디지털 정보자원의 지속적인 가치 및 진본성에 대한 평가 및 장기보존에 대한 실용성을 기준으로 큐레이션 대상 정보자원을 선별해야 한다. 보존 및 큐레이션에 드는 비용이 많기 때문에 선정 기준을 명시하며, 현재 및 미래의 이해관계자들이 접근 및 공유하기 위한 대상 자료인지 평가할

때 대상 자원의 생성 및 활용에 관한 업무 프로세스 확인과 분석 과정을 거쳐야 한다. 이용자 요구를 파악하고, 보존전략을 수립하고, 이를 통한 평가와 선정 정책은 정확한 데이터가 지속적으로 유지되도록 지원하는 필수요소이므로 성문화되어야 한다. 또한 디지털 자원의 평가 정책에는 재평가 원칙과 재평가 스케줄도 포함되어야 한다.

(4) 저장과 데이터 관리

리포지터리 정책과 표준 개발 및 유지를 위한 정책을 명시하고, 디지털 리포지터리의 감사와 인증절차를 평가할 수 있는 체크리스트가 제시되어야 한다. 최소한의 기준을 준수하는 양질의 데이터를 생성하고, 저장 기간 및 사용에 대한 기간이 설정되어야 하며, 리포지터리를 유지하고 연구결과물에 대한 저장, 혹은 선택되지 않을 경우에는 포기서류를 작성해야 한다. 기술과 표현정보에 대해서는 보존과 기술 정보를 제공하고 메타데이터의 스키마와 표준을 명시한다.

(5) 데이터 접근, 공유 및 재사용

데이터를 어떻게 이용 가능하게 하고, 사용 권한 및 제한에 적용하는 방법에 대한 설명과 접근 권한에 적용하는 과정과 사용 플랫폼에 대해 명시한다. 플랫폼의 선택은 기관의 목표와 재사용 기능, 또는 외부 이용자에게 제공하느냐에 따라 달라진다. 데이터 공유에 관한 사항은 제도적으로 기관의 데이터 보안 정책과 일치해야 하며, 데이터 및 연구결과물에 대해 리포지터리와 데이터 센터를 통해 제한 없이 접속 가능해야 한다.

데이터 재사용을 위한 최소한의 질적 기준은 상호 운용성을 준수하는 메타데이터 표준과 디지털 자료의 이용과 재사용에 대한 법률적인 허용 범위 및 디지털 자료의 인증과 절차에 의해 접근제어가 가능하다는 사항을 명시한다. 또한 데이터 인용을 위한 표준이 있어야 한다.

(6) 변환

적합한 데이터 구조 또는 파일 포맷 선정, 메타데이터 기술을 통해 마이그레이션 과정을 성문화하고, 오리지널 비트스트림 유지, 엄격한 품질 체크, 절차, 이용 등을 명시한다. 큐레이트된 데이터는 후 분석(post-analysis) 결과의 확인, 향후 후속 실험의 기반으로, 분석의 축적을 위하여 사용되므로 마이그레이션, 에뮬레이션 등의 변환 관련 기술이 명시되어야 한다. 데이터 세트 복사본과 데이터를 설명하는 문서 및 메타데이터를 함께 제출하도록 시스템이 있어야 하고, 데이터가 규칙적으로 백업되는지도 파악되어야 한다. 또한 관련 소프트웨어와 원본 데이터는 함께 유지되어야 한다. 다만 연구결과물이 CD-ROM, 비디오, 사진 등과 같이 영구적이지 않을 경우에 복사본을 제출한다.

라. 지원 규정

(1) 교육

디지털 큐레이션을 위해 큐레이터 양성 및 기존의 업무를 수행하고 있는 직원에게 교육을 통하여 디지털 정보자원 큐레이션을 위한 교육을 정책적으로 지원해야 한다. 또한 교육을 통하여 관련 이해관계자들이 다른 기관과의 협력을 위한 플랫폼을 제시할 수 있어야 한다.

(2) 홍보 및 가이드

이해관계자에게 큐레이션 관련 홍보 및 가이드를 위한 지침을 제공한다. 디지털 보존 및 큐레이션, 메타데이터에 적합한 파일 포맷 등의 자세한 내용을 다루고, 웹페이지를 통한 홍보 및 가이드 정책을 통하여 후원서비스가 이루어져야 한다.

(3) 모니터링

디지털 큐레이션을 위한 모니터링이 계속적으로 이루어져 생명주기 전반에 걸쳐 정책이 제대로 수행되고 있는지의 여부를 파악하고 반영한다.

이상과 같이 국내에서 적용이 가능한 큐레이션 정책 수립에 필요한 원칙 및 지침, 정책 수립, 생명주기 기반 큐레이션 정책, 지원 규정은 디지털 큐레이션 정책을 구현하기 위하여 최소한 필요한 항목이므로 실행을 위한 지원 인프라도 마련되어야 할 것이다. 또한 모니터링 메커니즘을 구현하고, 부적절한 조항은 개정하여 지속적인 큐레이션 정책이 이루어질 수 있어야 한다.

IV. 결론 및 제언

디지털 보존과 접근 및 재사용을 위한 디지털 큐레이션 정책은 현재 또는 미래의 이용을 위해서 필수적인 과정이 되었다. 이를 위하여 디지털 정보자원을 단순히 보존하는 것만이 아니라 가치를 부여하여 미래의 연구에 도움이 되는 기반을 마련하여 비용 및 시간을 절약하고, 공유할 수 있는 디지털 큐레이션 생명주기 단계별 정책 프레임워크를 개발하였다.

국내에서는 이경남·이소연이 전자기록을 중심으로 장기적 보존과 관리를 위한 아젠다를 법제 및 정책, 연구, 기술, 파트너십의 4개 영역으로 나누어 기관별 유형에 따른 정책 수립을 제안하였고, 이윤주·이소연은 진본 전자기록의 장기보존을 위한 정책 프레임워크도 제시한바 있다. 그러나 이들의 제안은 아카이빙 단계까지를 주체로 한 정책 연구이므로 디지털 큐레이션 정책을 수행하는 과정에서는 많은 보완점을 필요로 한다. 한편 Michael은 영국의 주요 연구재단의 정책, 실무 및

견해에 대해 연구하면서 큐레이션 대상으로 저널, 모노그래프, 컨퍼런스 자료, 논문, 데이터, 회색문헌 등의 연구 산출물의 충분한 활용을 위해 일관된 정책 프레임워크 개발이 필요하다고 제안하였다. 그러나 Michael의 디지털 큐레이션 정책 프레임워크 연구보고 또한 생명주기를 적용한 사례는 아니었다.

본고에서는 디지털 큐레이션 생명주기 단계별로 정책 프레임워크를 설계하였으며, 각 단계에 따른 지침을 제시함으로써 디지털 정보자원의 장기보존을 위한 실제적인 업무를 수행함에 있어 효율적으로 운영할 수 있는 기반을 마련하였고, 큐레이션을 위한 교육, 홍보 및 가이드, 모니터링 등에 대한 요구도 정책 개발에 포괄적으로 반영하였다. 앞으로 국가 차원에서나 각 기관에서 큐레이션 정책을 개발하기 위한 모델 프레임워크가 되기를 기대한다.

〈참고문헌은 각주로 대신함〉