

데이터 큐레이터의 핵심 직무 요건 고찰에 관한 연구*

An Investigation on Core Competencies of Data Curator

이 유 경 (You-Kyoung Lee)**

정 은 경 (EunKyung Chung)***

초 록

디지털 정보기술과 인터넷의 발달로 인해 데이터는 여러 학문과 사회 분야에서 의미 있는 결과를 도출하고 의사결정과 정책을 수립하는 데 있어서 중요하다. 따라서 대규모의 데이터를 추적하고 관리하는 업무를 담당하는 데이터 큐레이터는 데이터 관리의 효율성과 활용성을 높이는 데 필수적인 역할을 수행한다. 본 연구는 데이터 큐레이터의 핵심 직무 요건을 규명하고자 한다. 이를 위해 두 종류의 데이터를 수집하여 분석하였다. 첫째는 2011년부터 2014년까지 총 255건의 국외 데이터 큐레이터 구인 공고를 수집하였으며, 둘째는 국내 데이터 관리 실무자 5명과의 심층 면담을 실시했다. 수집된 두 가지 데이터는 선행연구에서 제시한 7가지 핵심 직무 요건을 바탕으로 범주화하여 분석하였다. 데이터 분석 결과 데이터 큐레이터의 핵심 직무 요건을 4가지 범주로 구분할 수 있다. 본 연구의 결과 데이터 큐레이터의 핵심 직무 요건은 효율적인 데이터 관리 직무 수행을 위해 이해관계자와 원활하게 소통할 수 있어야 하며, 이와 함께 정보기술 지식을 기반으로 데이터 관리 시스템을 구축하고 운영할 역량이 필요하다. 또한 데이터 관리 지식 및 전략을 이해하고 마지막으로 데이터 관리 이용자 교육 및 데이터 관련 서비스와 도구를 제공할 수 있어야 한다. 이를 통해 데이터 관리에 있어 4가지 범주로 구분되는 다양한 측면의 요건이 요구되고 있음을 도출하였으며, 이를 위한 데이터 관리에 대한 통합된 전문적인 교육과정의 필요성을 제안했다.

ABSTRACT

As the digital technologies and internet have advanced, data have centered in the process of meaningful scientific ramifications and policy making in a wide variety of fields. Data curator in charge of managing data plays a significant role in terms of improving the effectiveness and efficiency of data management and re-uses. The purpose of this study is to identify the core competencies for data curator. For achieving the purpose of this study, two sets of data were collected. First, a total of 255 job descriptions were collected from the web sites including ARL, Digital Curation Exchange, Code4lib, ASIS&T JobLine for the period of 2011-2014. Second, in-depth interviews with five data curators from four diverse organizations were collected. The two sets of data were analyzed into seven categories identified from the related studies. Findings of this study showed that core competencies for data curator were identified into four categories, communication skills, data management techniques, knowledge and strategies for data management, and instructions and service provisions for users. The implications of this study can be considered as integrated and professional curriculum developments for data curator with core competencies.

키워드: 데이터 큐레이터, 데이터 큐레이션, 디지털 큐레이션, 직무 요건, 데이터 관리 실무

Data Curator, Data Curation, Digital Curation, Job Requirement, Data Management

-
- * 본 연구는 이화여자대학교 일반대학원 석사학위논문 『데이터 큐레이터 직무 요건 분석 연구』의 내용을 기반으로 함.
- ** 이화여자대학교 일반대학원 문헌정보학(virginialay@naver.com) (제1저자)
- *** 이화여자대학교 사회과학대학 문헌정보전공 부교수(echung@ewha.ac.kr) (교신저자)
- 논문접수일자 : 2015년 8월 17일 논문심사일자 : 2015년 8월 28일 게재확정일자 : 2015년 8월 31일
- 한국비블리아학회지, 26(3): 129-150, 2015. [http://dx.doi.org/10.14699/kbiblia.2015.26.3.129]

1. 서론

인터넷과 정보기술의 발전으로 인해 데이터는 다양한 학문 분야에서 연구 수행 과정의 중요한 요소일 뿐만 아니라 여러 사회 분야의 문제 해결의 중심에 있다. 데이터 큐레이션은 디지털 형식으로 생산된 데이터를 생애 주기에 맞춰 저장·관리·보존하여 재사용을 도모하는 활동으로 정의할 수 있다(Tammaro, Ross, and Casarosa 2014). 데이터의 유형은 실험 결과, 관찰 결과, 통계 조사 자료, 재고 조사 결과물 등이며 분야별, 수집 매체별로 다양하다. 데이터는 연구 결과를 뒷받침할 수 있으며, 장기간 보존을 통해 통시적 메타 연구 및 학제 간 연구를 가능하게 하고 기존 연구에 새로운 가치를 부여한다(Dekker 2006).

따라서 최근의 연구 환경에서는 데이터가 중요한 역할을 수행하며, 대규모 데이터의 전문적 관리가 점점 더 중요한 상황이다. 이와 관련해 국외에서는 증가하는 디지털 데이터를 효과적으로 관리하기 위해 국가 단위를 비롯해 각 분야별 데이터 관리 기관을 중심으로 보존 활동이 이루어지고 있다. 또한 2000년대 들어서 오픈 데이터 정책을 추진해 정부 및 공공 기관에서 보유한 데이터를 공개하고 공유를 촉진하기 위한 노력을 하고 있다. 이 같은 맥락에서 데이터의 원활한 재활용을 위해 상호운용성이 높은 고품질의 데이터를 생산하는 것이 중요한 논의사항으로 부상했다. 이러한 관심은 정보 산업 및 연구 커뮤니티와 교육 영역에서 전문적인 데이터 큐레이터의 양성과 해당 직무를 정의하고자 하는 노력으로 이어졌다. 효과적인 데이터 큐레이션을 수행하려면 데이터가 생산되는 기관 내

부 및 외부 환경과 데이터 생애 주기에 대한 명확한 이해를 가진 실무자인 데이터 큐레이터의 역할이 중요하다.

데이터 큐레이터는 데이터 생산자와 함께 데이터 큐레이션을 수행하는 전문가를 말한다. 대규모 데이터를 선별·보존하며, 재사용이 가능하도록 관리한다. 또한 연구자와 정보기술 전문가를 연결하는 다리 역할을 수행한다(Lord et al. 2004; Stanton et al. 2011). 그러나 아직까지는 데이터 공유 및 재사용을 위한 일관된 큐레이션 활동이 이루어지기에는 실무 종사자들 사이에 전공 분야, 유입 경로, 경력 경로 등에서 많은 차이가 존재한다(Corrall 2012). 또한 Parkinson(2008)은 역량 있는 데이터 큐레이터를 식별할 수 있는 기준이 없음을 문제로 지적했다. 이와 함께 데이터 큐레이터는 발전 단계를 거치고 있는 새로운 직무 영역이므로 해당 직무에 관련된 용어가 가변적이며(Kim, Warga, and Moen 2013), 데이터 큐레이터의 직무를 단순하게 정의하기도 쉽지 않다는 연구 결과가 도출되고 있다. 데이터 큐레이터는 다양한 학문 분야, 기관 내부, 기관을 둘러싼 외부 조직, 데이터 관련 전문가들과의 협력이 필수적이기 때문이다. 이에 본 연구는 데이터 큐레이터가 직무를 수행하는데 있어 요구되는 핵심적인 직무 요건을 규명하려는 연구 목적을 가지고, 데이터 큐레이션 실무를 파악할 수 있는 구인 공고와 실무자 면담을 통해 현장에서 요구되는 직무 요건을 도출하고자 했다.

현 시점에서는 데이터 큐레이션이라는 용어가 독립적으로 정의되고 사용되지만, 동시에 여러 연구에서 디지털 큐레이션 개념의 세부 기능으로 포함되거나 혼용(Digital Curation Centre

2015: Gold 2010)되고 있다. 본 연구에서는 데이터 큐레이션을 '데이터를 체계적으로 보존해서 재사용을 촉진하고 그것에 가치를 부여하는 활동을 포괄하는 개념'으로 정의하고자 한다. 이와 함께 데이터 큐레이션을 수행하는 실무자를 '데이터 큐레이터'로 지칭하고자 한다.

2. 관련 연구

데이터 큐레이션은 그 용어에 대한 정의와 연구가 최근 들어 증가하는 추세이며, 본 연구와 관련된 연구는 크게 두 그룹으로 구분될 수 있다. 첫째는 데이터 큐레이션의 효용성에 대한 논의를 다루는 연구이며, 두 번째는 데이터 큐레이터 직무를 규명하는 연구이다.

2.1 데이터 큐레이션의 효용성

데이터 큐레이션의 효용성 관점에서 이루어진 연구는 데이터 큐레이션을 사서의 새로운 역할 및 대학 내 기관 리포지터리의 확장된 서비스 역할로 인식했다. 이를 기반으로 e리서치를 비롯한 학제 간 연구 환경 속에서 대학도서관과 연구도서관의 역할과 효용을 도출하고자 하는 연구들이 진행되고 있다. Martinez-Uribe와 MacDonald(2008)는 도서관과 사서에게 데이터 큐레이션은 기회이자 도전이라고 주장했다. e리서치 환경에서는 데이터의 공유가 중요한 논의사항이다. 학문 영역별로 데이터를 관리하고 기술하며 저작권을 부여하는 일련의 공유 및 출판 과정을 체계적으로 수행할 필요가 있다. 그러나 특정 학문 분야의 전문성만으로

는 효과적인 데이터 관리가 어려운 점을 지적했다. 따라서 데이터 조직 및 관리에 대한 오랜 경험과 지식을 축적한 도서관과 사서의 역할이 e리서치 환경에서 강점으로 작용할 수 있다고 주장했다. Gold(2010)는 국가적인 데이터 큐레이션 인프라 구축에 있어 도서관의 역할이 중요함을 강조했다. 단기적으로는 3개의 계층으로 이루어진 역할 모델을 제시했다. 즉, 국가 차원의 인프라 구축, 대학 차원의 인프라 구축, 전문 인력 양성이다. 이와 함께 장기적 관점으로는 다음과 같이 4가지 차원에서 제시했다. 데이터 큐레이션 실무자 커뮤니티 조성, 시스템이 아닌 서비스 개념으로 접근, 기관 리포지터리 활성화, 이를 통한 상호운용성 제고와 주제 영역을 넘나드는 파트너십 구축이다. Heidorn(2011)은 도서관이 오랫동안 수행해 온 정보와 지식의 큐레이션 기능에 주목했다. 증가하는 데이터, 정부 연구비 지원 기관의 데이터 관리 정책, 학제 간 연구 등의 사례를 살펴보고, 디지털 데이터 큐레이션이 가능한 최적의 기관인 도서관이 데이터 관리 과정에서 적극적인 역할을 해야 한다고 주장했다. 단지 연구도서관뿐만 아니라 모든 유형의 도서관에서 데이터 큐레이션 역할이 요구되고 있음을 강조했다. Kruse와 Thestrup(2014)은 도서관이 연구데이터를 관리하는 데 장애요소가 무엇인지 조사했다. 그 결과 연구데이터 관리에 대한 공통적인 정보기술 인프라가 없고, 데이터의 규모 및 데이터 포맷, 연구 환경의 차이 등을 이유로 대학도서관들이 소극적인 전략을 취하고 있음이 드러났다. 해당 연구에서는 대학 및 연구기관이 개별적으로 어떠한 전략을 수행하는지 역시 조사·분석했다. 이 과정에서 덴마크주립도서관이 참여한

데이터 관리 및 큐레이션 프로젝트 중 LARM, Netachive, Mediestream, DigHumLab에 주목했다. 위의 프로젝트는 라디오 데이터베이스, 덴마크의 동적 인터넷, 문화유산 자원 디지털화, 디지털인문학 자원 관리 교육 및 지식 이전을 목적으로 수행되었다. 위의 프로젝트 분석을 통해 디지털 자원 보존 및 연구데이터 관리를 위해 도서관이 할 수 있는 다양한 활동을 제시했다.

2.2 데이터 큐레이터의 직무 분석

데이터 관리를 수행하는 데이터 큐레이터는 효율적인 데이터 큐레이션의 중요한 요소 중의 하나로 볼 수 있다. 이에 데이터 큐레이터의 직무 요건을 파악하려는 연구가 수행되었다. 이들 연구는 도서관의 데이터 서비스 관련 사서들 또는 각 주제 분야 데이터 큐레이션 실무자를 대상으로 한 설문 조사나 구인 공고 분석에 토대를 두고 있다. Stanton et al.(2011)은 e사이언스 환경에서 연구자와 IT 엔지니어 사이에서 발생하는 대규모 정보의 관리 문제들을 다루었다. 이를 해결하는 방법으로 사서의 역할이 e사이언스 전문가로 확장될 수 있다고 판단했다. 이를 입증하기 위해서 연구진은 데이터 관리 실무자로 구성된 포커스 그룹과 데이터 관리 인턴십에 참여했던 대학원생을 대상으로 면담을 실시했다. 면담은 Cronshaw와 Fine가 '기능적 직무 분석'에서 제시한 데이터(data), 사람(people), 도구(thing)의 3가지 범주를 바탕으로 e사이언스 전문가의 직무를 분류했다. 해당 연구에서는 e사이언스 전문가의 핵심 기능을 데이터 관리, 기술 관련 스태프와 기술과

관련 없는 스태프 사이의 원활한 소통, 연구 관련 기능으로 제시했다. Creamer et al.(2012)은 데이터 큐레이션 수행을 위해 요구되는 핵심 역량을 도출하고자 미국 뉴잉글랜드 지역의 보건학 및 과학도서관 사서를 대상으로 설문 조사를 실시했다. 응답한 사서들은 연구자들에게 데이터 큐레이션과 관리 서비스를 적극 홍보하고 지원할 의지는 있지만, 큐레이션과 관련된 전문 기술(technology)의 추가적인 교육을 필요로 했다. 특히 데이터 큐레이션 관련 기술은 한 사람의 전문가가 보유하고 수행할 수 없으며 팀의 형태로 수행해야 한다는 의견이 많았다. 응답자들은 데이터 관련 지적재산권 이슈나 데이터 문해(文解)와 관련한 주제에 많은 관심을 보였으며, 무엇보다도 기관의 투자가 중요하다고 응답했다. Kim, Warga와 Moen(2013)은 도서관의 업무 영역이 데이터 관리에까지 확장됨에 따라 그것을 담당할 정보 전문가가 필요하고, 이들을 양성할 교육 과정에 대해 연구가 필요하다고 판단해 도서관 및 디지털 큐레이션 분야 구인 공고를 분석했다. 북미도서관협회, 미국연구도서관협회, 전문도서관협회, Digital Curation Exchange(이하 DCE)의 구인 공고를 수집해 직책, 기관, 학위, 경력, 요구되는 지식·능력, 수행해야 할 직무에 초점을 맞추었다. 분석 결과, 선별, 입수, 보존 등 디지털 큐레이션 단계별로 수행하는 기관이 분산되어 있었다. 또한 디지털 큐레이션에 대한 명확한 업무 정의가 부족했다. 따라서 디지털 큐레이션 실무는 디지털 큐레이션 전 과정에 이르는 정책 및 절차에 대한 이해가 바탕이 되어야 한다고 주장했다. 이를 토대로 업무 기술(skill)과 기술적 역량 모두를 아우르는

교육이 이루어져야 한다고 했다. 또한 데이터가 생산되는 주제 분야 연구자에게 리에종 서비스, 교육 등 전문화된 연구 컨설팅 교육을 제공해야 한다고 주장했다. Tammaro, Ross와 Casarosa(2014)는 문화유산 기관에서 데이터 큐레이터에 대한 요구가 증가하고 있으며, 따라서 데이터 큐레이션 전문가를 양성하는 교육에 관심이 증대되고 있음에 초점을 맞추었다. 이에 1) 사서, 아키비스트, 학예사의 공통적인 직무 역량이 있는지, 2) 각 영역별 데이터 전문가가 되기 위해서는 각기 다른 지식이 필요한지, 3) 기존의 사서, 아키비스트, 학예사의 직무 요건과 주제 분야 연구에서 요구되는 데이터 큐레이터의 역할 및 기술이 동일한지의 세 가지 연구 질문을 도출하고 이를 검증하고자 했다. 연구진은 유관 기관, 이용자 서비스, 데이터 관리, 기술 분야의 4가지 영역에서 데이터 큐레이션 수행에 필요한 직무 역량을 도출했다. 그리고 데이터 큐레이션 전문가 양성 교육 프로그램에 대한 이론적 체계와 구조가 다양한 주제 분야에 맞출 수 있도록 유연하게 설계될 필요가 있다고 주장했다.

국내에서도 디지털 큐레이션에 대한 정책 및 관리 체계 구축 방안과 관련된 연구(김선대 외 2010; 안영희 2011; 박승진 외 2011; 신영란 2012)와 디지털 큐레이션 전략을 적용해 대학 도서관 서비스 개선 방안(안영희, 박옥화 2009; 김지현 2014; 한나은 2014)을 도출하는 연구도 이어지고 있다. 안영희와 박옥화, 신영란, 김지현의 연구에서는 안정적인 체계적인 디지털 큐레이션 서비스와 체계를 구축하기 위해서는 역량 있는 실무자 양성이 요구된다고 강조했다. 그러나 국내에서는 국내 환경에 적합한

데이터 큐레이션을 수행하는 전문 인력의 직무 역량을 도출한 연구가 미비한 실정이다. 따라서 본 연구는 데이터 큐레이션과 관련된 국외 구인 공고와 국내 데이터 관리 실무자 심층 면담을 통해 데이터 큐레이터의 핵심 직무 요건을 규명하고자 한다.

3. 연구 설계

3.1 데이터 수집

데이터 큐레이터의 핵심 직무역량을 규명하기 위한 연구 목적을 달성하기 위해서 본 연구는 두 종류의 데이터를 수집하였다. 첫 번째로 데이터 큐레이션 실무와 채용 공고가 상대적으로 활발하게 이루어지는 국외의 구인 공고를 수집했다. 구인 공고 데이터 수집은 선행연구에서 참조했던 북미연구도서관협회(Association of Research Libraries), 데이터 큐레이션 직업 정보 교환이 발생하는 Digital Curation Exchange와 추가로 Code4lib, 정보과학기술협회(ASIS&T)의 JobLine을 선정했다. 수집 대상은 2011년 1월 1일부터 2014년 12월 31일까지 최근 4년간 게시된 구인 공고이고, 중복 공고를 추려내어 최종적으로 255건의 데이터를 수집했다. 두 번째로 국내 데이터 관리 실무자 5명과 심층 면담을 통해 데이터를 수집했다. 면담 대상은 국내 대규모 데이터 관리 전략으로 디지털 큐레이션 적용을 제안한 신영란(2012)의 연구에서 국내 인문사회 분야 데이터 아카이브로 선정한 20세기민중생활사연구소와 한국사회과학데이터센터, 소속 정보 비공개를 요청한 인문사회

〈표 1〉 데이터 큐레이션 실무자 면담지 구성

항목	세부 항목	문항수
피면담자 일반 현황	<ul style="list-style-type: none"> • 연령대 • 전공 분야/최종 학위 • 소속 기관 유형 • 관리하는 데이터 유형 • 유관 경력 유무 • 유관 경력 기관 및 기간 	6
데이터 큐레이션 인식 정도	<ul style="list-style-type: none"> • 업무 인식 • 관련 교육 	2
데이터 큐레이션 직무 요건	<ul style="list-style-type: none"> • 이해관계자 정의 및 구성 • 큐레이션 관련 기술 활용 (2 문항) • 의사소통 및 대인관계 • 큐레이션 및 보존 전략 수행 • 계획·관리·평가 활동 • 최신 동향 분석 • 시스템·모델 및 모델링 • 서비스 제공 	9

분야 데이터 아카이브를, 박승진 외(2012)의 연구에서 과학기술 분야 데이터 아카이브로 정의한 KISTI를 선정했다. KISTI의 경우, 과학 데이터 전략 담당자에게 전화와 서면으로 문의했고, 문의 결과 논문의 주제에 더욱 적합한 한국지질자원연구원 실무자를 소개받을 수 있었다. 그리하여 총 4개 기관, 5명의 실무자를 대상으로 심층 면담을 실시했다. 면담지 구성은 3개 항목과 17개 문항으로 구성했다. 실무자 면담지는 데이터 큐레이터 직무 요건을 연구한 Stanton et al.(2011), Creamer et al.(2012), Kim et al.(2013)의 연구를 바탕으로 반구조화된 형식으로 구성했다. 피면담자 일반 현황, 데이터 큐레이션 인식 정도, 데이터 큐레이션 직무 요건으로 3개 항목을 구성하고 각 항목별로 세부 항목을 부여했다(〈표 1〉 참조). 데이터 큐레이션 직무 요건 항목은 구인 공고 분석 항목으로 참조한 Kim et al.(2013)의 7가지 직무 요건을 기

준으로 실제 수행하는 업무에 관해 질문했다. 동의하지 않은 1명의 피면담자를 제외한 4명의 피면담자와 녹음을 했고, 녹음된 파일을 전사했다.

3.2 데이터 분석

수집된 두 종류의 데이터는 일반 현황 분석과 직무 요건 분석으로 구분되었다. 첫 번째 데이터인 구인 공고는 일반 현황 분석과 직무 요건 분석으로 나누어 범주화를 수행했다. 첫째, 구인 공고의 일반 현황을 살펴보기 위한 분석 항목은 수집한 구인 공고의 구성을 바탕으로 Choi와 Rasmussen(2009)과 Kim et al.(2013)의 연구의 범주화를 사용했다(〈표 2〉 참조).

둘째, 구인 공고에 나타난 데이터 큐레이터 직무요건은 Kim et al.(2013)가 제시한 7가지 항목을 기준으로 분석했다. 7가지 항목은 〈표 3〉에

〈표 2〉 데이터 큐레이터 구인 공고 일반 현황 분석 항목 구성

항 목	항목 예시	
직무명	서비스, 정보기술, 데이터 유형, 기타	
기관 유형	대학도서관 및 연구도서관, 국가도서관, 공공도서관, 정부기관	
경력 사항	유관 기관 경력	대학도서관 및 연구도서관, 아카이브, 박물관, 영리기관, 경력 기관 명시 없음
	경력 기간	1~2년, 3~5년, 5년 이상, 요구 있으나 기간 명시 없음, 요구 없음

〈표 3〉 직무 요건 정의와 세부 항목 구성

항 목	정의	세부 항목	
큐레이션 관련 정보기술 활용 요건	데이터 큐레이션 활동을 지원할 수 있는 정보기술을 활용해서 데이터 관리 시스템을 개발하고, 관련 도구와 응용프로그램을 식별하고 활용할 수 있는 요건	정보기술 개발 및 운영 요건	<ul style="list-style-type: none"> •웹기반 마크업 언어 •데이터베이스 관리 시스템 언어 •스크립팅 및 프로그래밍 언어 •문서 작성 및 그래픽 도구 활용 •운영 체제 활용
		정보기술 도구 활용 요건	<ul style="list-style-type: none"> •리포지터리 플랫폼 •오픈소스 소프트웨어 활용 •기록관리 시스템 및 콘텐츠 이관 도구 •도서관 통합관리시스템 •데이터 분석 도구 활용
의사소통 및 대인관계 요건	다양한 이해관계자들과 효율적으로 소통하는 요건	<ul style="list-style-type: none"> •구술 및 기술 소통 •독립적 업무 및 협업 •리더십 	
큐레이션 및 보존 활동 요건	디지털 큐레이션 생애 주기 모델에서 정의한 일련의 활동을 이해하고 수행하는 요건	<ul style="list-style-type: none"> •큐레이션 전략 및 절차 •메타데이터 표준 •저장소 아키텍처 •보존 및 기타 표준 •다양한 매체 유형 관리 •학술연구의 정책 및 절차 •저작권 및 개인 정보 관리 •검색 관련 기술 	
계획·관리· 평가 요건	디지털 큐레이션과 관련된 프로젝트, 프로그램, 서비스를 계획하고 조정하고 실행하고 평가하는 요건	<ul style="list-style-type: none"> •프로젝트, 프로그램, 서비스의 계획·수행·관리·평가 •직원 감독 •연구과제 신청서 작성 	
최신 동향 분석 요건	디지털 큐레이션 최신 동향·기술·실무를 식별하고 자원을 활용하는 요건	<ul style="list-style-type: none"> •최신 연구 및 기술 동향 분석 •전문성 개발 	
시스템·모델 및 모델링 요건	디지털 큐레이션과 관련된 시스템과 업무 흐름, 개념 모델에 대해 비판적으로 분석하는 요건	<ul style="list-style-type: none"> •분석적 사고 	
서비스 요건	커뮤니티에서 디지털 큐레이션과 관련한 요구를 식별하고 대응하는 서비스를 제공하는 요건	<ul style="list-style-type: none"> •리에종 및 서비스 지원 •교육 및 컨설팅 •홍보 활동 	

서 살펴볼 수 있는 바와 같이, 1) 큐레이션 관련 기술 활용 요건, 2) 의사소통 및 대인관계 요건, 3) 큐레이션 및 보존 활동 요건, 4) 계획·관리·평가 요건, 5) 최신 동향 분석 요건, 6) 시스템·모델 및 모델링 요건, 7) 서비스 요건이다. 이러한 7가지 항목과 각각의 항목에 포함되는 세부 항목을 이용하여 분석했다.

4. 결 과

4.1 국외 구인 공고 분석

국외 구인 공고에 나타난 일반 사항은 직무명, 기관 종류, 경력 사항으로 구분하여 살펴볼 수 있다. 우선, 총 255건의 구인 공고에 나타난 데이터 큐레이터 관련 직무명은 8건의 중복 데이터를 포함해 총 263건이며, 분석 결과 직무명은 두 개 이상의 용어로 조합된 다양한 직무명을 찾아볼 수 있다. 본 연구에서는 서비스 관련 직무명, 정보기술 관련 직무명, 데이터 유형 관련 직무명, 기타 직무명의 4가지 세부 항목으로 범주화하고 복수의 항목에 해당하는 8건의 경우는 각 세부 항목에 중복으로 분류했다. 분석 결과를 보면 서비스 관련 직무명이 전체의 50%(131건)로 가장 높은 빈도를 보였다. 정보기술 관련 직무명이 21%(55건)로 두 번째로 높게 나타났으며, 데이터 유형 관련 직무명 15%(40건), 기타 직무명 14%(37건)였다. 이 중 정보기술 관련 직무명은 4년 동안 꾸준한 증가세를 보이고 있었다. 두 번째, 데이터 큐레이터 구인 공고를 낸 기관의 종류를 분석한 결과 총 9종의 기관으로 나타나났다. 총 255건 중에

서 가장 높은 빈도로 관련 구인 공고를 게시한 기관은 대학도서관 및 연구도서관으로, 비중이 가장 큰 69%(176건)로 나타났다. 이 밖에도 영리 기관 7%(19건), 비영리 기관 7%(17건), 아카이브 6%(15건), 국가도서관 4%(9건), 정부 기관 3%(8건), 박물관 2%(6건), 공공도서관 2%(4건)가 그 뒤를 이었다. 데이터 아카이브는 0%(1건)로 가장 적은 수를 나타냈다. 세 번째는 구인 공고에 나타난 필요한 유관 기관의 경력을 분석한 결과, 대학도서관 및 연구도서관 업무 경력을 가장 선호하는 것으로 나타났다. 10건의 중복 분류를 포함한 265건을 대상으로 분석한 유관 기관 경력 분석 중 유관 경력 기관을 요구한 공고는 39%(102건)이고, 명시하지 않은 공고는 61%(163건)이다. 유관 경력 기관을 요구한 공고 39%(102건) 중 대학도서관 및 연구도서관 업무 경력을 요구한 구인 공고는 총 24%(63건)로 가장 높은 선호도를 나타냈다. 유관 경력 기관을 명시한 공고를 100%로 환산하면 대학도서관 및 연구도서관 경력 요구는 그중 61%를 차지한다. 두 번째는 아카이브 업무 경력으로 10%(26건), 박물관 업무 경력은 4%(10건), 마지막으로 영리 기관의 업무 경력은 1%(3건)로 조사되었다. 2012년부터는 박물관과 영리 기관의 업무 경력에 대한 요구가 소폭이지만 증가 추세를 보이고 있었다. 유관 경력 기간 요구는 총 255건의 구인 공고 중 54%(138건)가 발생했으며, 경력 기간을 요구하지 않은 공고는 46%(117건)이었다. 경력 기간을 요구한 공고 중에서 가장 선호되는 경력 기간은 3~5년(27%)으로 나타났다. 3~5년의 경력 기간을 요구한 구인 공고는 2011년 39%(7건), 2012년 27%(22건), 2013년 18%(15건), 2014년

34%(25건)이다. 이중 1~2년의 경력 기간 요구가 2011년은 5%(1건)였으나, 2012년 14%(11건), 2013년 15%(12건), 2014년은 12%(9건)로 점차 증가하는 추세를 보였다.

데이터 큐레이터 구인 공고에 나타난 Kim et al.(2013)의 7가지 전체 직무 요건 출현 빈도를 분석한 결과는 <표 4>와 같다. 큐레이션 관련 정보기술 활용 요건이 가장 높은 20%(213건)으로 나타났고, 두 번째로 의사소통 및 대인관계 요건이 19%(200건)로 나타났다. 세 번째로 큐레이션 및 보존 활동 요건 18%(196건), 네 번째로 계획·관리·평가 요건 15%(161건), 다섯 번째로 최신 동향 분석 요건 11%(112건), 여섯 번째로 시스템·모델 및 모델링 요건 10%(103건), 일곱 번째로 서비스 요건이 7%(79건)로 조사되었다. 각 연도별로 출현 빈도를 살펴보면, 2011년은 의사소통 및 대인관계 요건이 24%(17건)로 가장 높은 빈도를 나타냈지만, 2012년, 2013년, 2014년은 모두 큐레이션 관련 정보기술 활용 요건이 각각 20%, 20%, 19%로 가장 높은 빈도를 차지해 큐레이션 관련 기술

활용에 대한 요구가 증가하고 있음을 알 수 있다. 더불어 의사소통 및 대인관계 요건, 큐레이션 및 보존 활동 요건, 계획·관리·평가 요건이 매년 높은 빈도로 출현하고 있다.

각 요건을 구성하고 있는 세부 항목의 출현 빈도를 살펴보면 다음과 같다. 큐레이션 관련 정보기술 활용 요건의 경우 정보기술 개발 및 운영 요건과 정보기술 도구 활용 요건 두 개의 하위 항목을 추가로 구성한 뒤 세부 항목을 부여했다. 정보기술 개발 및 운영 요건에서는 5개의 세부 항목 중에서 웹기반 마크업 언어, 데이터베이스 관리 시스템 언어, 스크립팅 및 프로그래밍 언어에 대한 요구가 각 세부 항목별 4년 합계 평균이 26%, 25%, 25%로 높은 빈도로 출현해 실무에 있어 그 요구가 높다는 것을 알 수 있었다(<표 5> 참조).

정보기술 도구 활용 요건 중에서는 리포지터리 플랫폼 활용에 대한 요구가 연도별 빈도와 4년 합계 모두 가장 높은 순위로 나타났다. 대학도서관을 비롯한 다양한 기관에서 기관 내 데이터 관리를 위해 리포지터리를 활용하는 능

<표 4> 데이터 큐레이터 구인 공고 직무 요건 빈도 분석

직무 요건 항목	2011		2012		2013		2014		합계	%
	건	%	건	%	건	%	건	%		
큐레이션 관련 정보기술 활용 요건	15	21	68	20	68	20	62	19	213	20
의사소통 및 대인관계 요건	17	24	56	17	65	19	62	19	200	19
큐레이션 및 보존 활동 요건	11	15	60	18	63	19	62	19	196	18
계획·관리·평가 요건	11	15	49	15	52	15	49	15	161	15
최신 동향 분석 요건	7	10	35	11	40	12	30	10	112	11
시스템·모델 및 모델링 요건	6	8	33	10	27	8	37	12	103	10
서비스 요건	5	7	30	9	24	7	20	6	79	7
계	72*	100	331*	100	339*	100	322*	100	1,064*	100

* 중복 출현 포함

〈표 5〉 큐레이션 관련 기술 활용 요건 빈도 분석 - 정보기술 개발 및 운영 요건

요건별 세부 항목	단위(건, %)									
	2011		2012		2013		2014		합계	%
	건	%	건	%	건	%	건	%		
웹기반 마크업 언어 지식 (예: XML, XSLT 등)	6	29	14	21	19	33	20	24	59	26
데이터베이스 관리 시스템 언어 지식 (예: MySQL, Oracle 등)	4	19	19	28	15	26	21	25	59	26
스크립팅 및 프로그래밍 언어 지식 (예: Java, PHP, Perl 등)	6	29	20	29	11	19	20	24	57	25
문서 작성 및 그래픽 도구 활용 (예: MS 오피스, 포토샵 등)	3	14	6	9	10	17	10	12	29	12
운영 체제 활용 (예: MS 윈도우즈, Linux 등)	2	9	9	13	3	5	12	15	26	11
계	21*	100	68*	100	58*	100	83*	100	230*	100

* 중복 출현 포함

력이 높은 비중으로 요구되고 있었으나 연도별로는 빈도가 감소하는 추세를 보였다. 반면 데이터 분석 도구 활용에 대한 요구가 증가하고 있음을 알 수 있었다. 2011년 0%(0건)에서 2012년 2%(1건), 2013년 15%(6건), 2014년 17%(9건)로 매년 증가세를 유지하며 2011년 대비 17% 증가했다. 데이터 이용자의 요구가 점차 세분화되고 다양해짐에 따라 데이터 분석 도구를 활용하고 서비스를 제공하는 활동이 중요하게 요구되고 있었다.

의사소통 및 대인관계 요건에서는 구술(oral) 및 기술(written) 소통 활동이 합계 평균 67%로 가장 높은 빈도로 출현했으며, 큐레이션 및 보존 활동 요건 중에서는 큐레이션 전략 및 절차 지식이 43%(290건)로 4년 합계 기준으로 가장 높은 빈도를 보였다. 큐레이션 전략 및 절차 지식은 연도별 빈도분석에서도 증가세를 나타냈다. 두 번째로 메타데이터 표준 지식이 2011년부터 2014년까지 매년 20%를 차지하며 고른

빈도로 출현하고 있다. 계획·관리·평가 요건은 데이터 큐레이션과 관련한 프로젝트, 프로그램, 서비스의 계획·수행·관리·평가하는 항목이 4년 합계 52%로 가장 선호되고 있음을 알 수 있다. 이와 함께 연구과제 신청서 작성에 대한 요구가 점차 증가세를 보이고 있었는데, 국외의 경우 주요 연구비지원기관에서 연구비를 지원받아 수행된 연구의 산출물 제출에 데이터 관리 계획(Data management plan)을 요구하는 정책을 펴고 있기 때문으로 해석할 수 있다(김지현 2014). 마지막으로 최신 연구 및 기술 동향 분석과 리에종 및 서비스 지원 활동이 높은 빈도로 요구되고 있었다.

4.2 국내 실무자 면담 분석

피면담자인 국내 데이터 큐레이션 실무자의 일반 현황은 연령대, 전공 분야와 최종 학위, 소속 기관 유형과 관리 데이터 유형, 유관 경력 여

부와 해당 기관 및 경력 기간으로 구분했다. 피면담자의 구성은 다음과 같다. 4개 기관 5명이 면담에 참여했으며 전공은 문헌정보학 2명, 컴퓨터 공학 2명, 지리공간데이터 1명이었다. 연령대는 30대가 2명, 40대가 2명, 60대 1명이다. 소속 기관은 정부기관 1명, 사단법인 1명, 대학교 1명, 대학교 연구소 1명 정부출연 연구기관 1명이다. 피면담자 중 2명은 현재 소속된 기관 이전에 대학도서관, 연구소 등에서 유관 경력이 있었다(〈표 6〉 참조).

데이터 큐레이션 업무의 인식에 대해서 피면담자들은 자신의 업무가 데이터를 수집하고 지속적으로 운영·관리하여, 구축된 데이터가 재사용될 수 있도록 촉진하는 활동이라고 정의하고 있었다. 그러나 시스템 구축 단계부터 데이터 큐레이션 전략을 적용했다기보다는, 시스템을 운영하고 데이터를 관리하는 실무 수행 과정에서 데이터 큐레이션 개념을 접한 사례가 대다수였다. 또한 피면담자 자신이 데이터 큐레이션에 대한 전문적인 교육을 받은 경험이 없다고 답하였다. 그러나 데이터 큐레이터 직무에 대한

개념과 해당 분야, 업무 범위 등을 정의하는 데 어려움이 있다고 응답해, 데이터 큐레이터의 필요성을 인지하고는 있지만 해당 직무를 명확하게 정의하는 것의 어려움을 드러내고 있었다.

“우리가 디지털 큐레이션을 안 것은 얼마 안 돼요. 그렇지만 우리가 당시에 디지털 큐레이션 역할을 한 거죠. 자료를 받고 관리하고, 어떻게 보여 줄 것인가를 고민하고... .”(P3-1, 60대, 경력 25년차)

“데이터 큐레이터가 옛날에 문헌(정보) 쪽에서 하는 사서의 일(로 봐야 할지? 아니면) 전문 분야별로 데이터의 성격이나... 품질관리까지 할 수 있는지... 굉장히 정의하기 어렵더라고요. 저희도 막연한 생각만 갖고 있어요. 외국 사례를 보면 이런 게 필요할 것 같은데... 사실 이런 업무를 하는 사람이 있는지도 궁금하고...”(P4, 40대, 경력 19년차)

이에 본 연구는 피면담자와 인터뷰를 커뮤니

〈표 6〉 피면담자의 일반 현황

피면담자	연령대	전공분야	최종 학위	소속기관 유형	관리데이터 유형	유관 경력		
						경력 여부	경력기관	경력 기간
P1	40대	문헌정보학	석사	교육부 산하 정부기관	디지털화된 이미지 데이터	○	대학도서관, 연구소	12년
P2	30대	문헌정보학	석사	교육부 산하 사단법인	조사 및 통계자료	X	-	8년
P3	P3-1	물리학, 컴퓨터공학	박사	대학교	멀티미디어, 텍스트	○	검색엔진 개발 연구소	25년
	P3-2	컴퓨터공학	석사	대학교 연구소	멀티미디어, 텍스트	X	-	8년
P4	40대	지리공간 데이터	박사	정부출연 연구기관	지리공간 데이터	X	-	19년

케이션, 정보기술, 계획·관리·평가 활동, 최신 동향 분석, 서비스 제공으로 구분하여 분석했다.

4.2.1 데이터 큐레이터와 커뮤니케이션

현실적으로 데이터 큐레이션 실무는 다양한 이해관계자들과 협업을 통해 이루어진다. 데이터 큐레이터 실제적인 업무에서 주로 의사소통을 수행하는 이해관계자는 '정보기술 관련 개발자와 유지보수자'인 것으로 나타났다. 정보기술 개발자와 유지보수자란 보존 시스템을 구축하고 관리하는 데 있어 기술적인 협력 관계를 맺게 되는 정보기술 전문가를 말한다. 기관 내 전산 관리자 및 데이터베이스 개발자, 웹서비스 개발자, 서버 관리자 등이 포함된다. 데이터 큐레이션 실무자는 정보기술 개발자와 긴밀한 협업 관계에 있다는 답변을 찾아볼 수 있다.

“자료를 수집팀에서 수집해 와서 직접 정리하기도 하지만, 시스템 담당자와 상의를 하게 되죠. 내용은 수집팀 담당자들이 더 잘 알고 있지만 전체적인 데이터베이스 구조에 대한 이해가 필요한 부분에 있어서 저와 상의하는 부분이 있고요.”(P1, 40대, 경력 12년차)

“데이터베이스 개발과 관련해서는 개발자들이랑 개발 업체랑 협업을 해야 하고... 개발자도 여러 종류가 있는데 서버 관리 하는 분, 웹 개발 하는 분, 모듈 쪽 개발해 주는 분들과 일해야 하고요.”(P2, 30대, 경력 8년차)

피면담자 모두 업무 활동 중에서 이해관계자

들과의 소통이 중요하고 빈번하게 일어난다고 답했다. 주로 기관에서 설정한 데이터 관리 방안 및 요구에 맞추어 실무를 진행될 수 있도록 소통하는 업무라고 응답했다. 특히 데이터 생산자와의 협업 과정에서는 데이터 관리와 시스템 활용에 대한 이해 정도에 차이가 존재해, 그 간극을 메우기 위한 소통 과정이 필요했다. 기관에서 통합적인 데이터 관리 사업을 추진하기 위해서는 주제 분야별 데이터 생산자들을 설득하고, 데이터 관리 시스템 구축 사업을 이해하고 참여할 수 있도록 공감대를 형성하는 업무가 이루어져야 한다고 답했다. 그러나 기관 내부의 이해관계자들과 공감대를 형성하는 것이 어렵다는 답변이 함께 제시되었다. 데이터를 공개하고 공유하는 문화가 정착되지 않은 상태에서, 실무자가 단독으로 통합 시스템을 구축하고 큐레이션 전략의 적용을 주장하는 것이 쉽지 않다는 것이다.

“데이터 생산 분야도 다양하고 사람도 다양하기 때문에 (커뮤니케이션이) 정말로 필요하죠. (데이터 통합 관리를 하려면) 사실은 전체적인 사람들의 인식이 굉장히 중요해요. 그거하고 협조라든지... 이런 것들이 굉장히 중요하기 때문에 (커뮤니케이션이) 필수적인 거죠.”(P4, 40대, 경력 19년)

“공감대 형성이 중요해요. 데이터 큐레이션 개념이 기관 내부에 없는 상황에서 이것을 업무에 적용하기 위해 관계자를 설득하거나 혹은 그런 설득 활동을 혼자 진행한다는 것은 정말 힘든 일이거든요.”(P1, 40대, 경력 12년차)

4.2.2 데이터 큐레이터와 정보기술

큐레이션 관련 정보기술에 대한 지식은 보존 시스템 구축 및 관리와 데이터 분석 도구 활용에 필요하다고 응답했다. 피면담자들은 데이터베이스 및 보존 시스템 구조에 대한 지식을 바탕으로, 데이터 생산 시점부터 보존에 필요한 데이터를 생산하는 데 직·간접적으로 참여한다고 답변했다. 직접적으로 참여한다고 답변한 피면담자의 경우, 기관 내부의 통합적인 데이터 보존 인프라를 구축한다고 했다. 구축 과정은 다음과 같았다. 데이터 수집 및 시스템 구축과 연관 있는 사업팀별 정보 요구를 파악하고, 그것을 기반으로 절차 및 워크플로우를 반영한 모듈을 설계한다. 이를 바탕으로 정보기술 개발자가 코딩을 하고, 완료가 되면 실제 이용자를 대상으로 시연을 한다. 이후 시연 과정에서 도출된 수정 내용을 반영해 구축을 완료한다고 답변했다. 구축이 완료된 뒤에도 데이터 수집 및 생산이 지속적으로 이루어지므로, 피면담자가 데이터 생산자들이 원활하게 데이터를 입력할 수 있도록 시스템을 관리하고 운영한다고 답변했다. 총괄 업무를 맡은 피면담자의 경우, 시스템 구축 및 운영에 간접적으로 참여하고 있었다. 실무자로서 데이터베이스 구축 및 관리 업무를 수행하던 시기에는 데이터 수집에 직접 참여했으나 총괄직무를 맡은 현재는 데이터 수집, 데이터베이스 구축 및 관리, 서비스 지원 업무까지 모두 관여한다고 답변했다.

데이터 분석 도구를 활용하는 업무와 관련해서는 기관의 소장 데이터를 재활용해 대규모 데이터에 존재하는 유용한 상관관계를 발견하려는 목적으로 데이터 분석 도구를 활용한다고 했다. 특히 분석 도구는 디지털 포맷으로 생산된

데이터가 대규모로 축적·관리되는 과학기술 분야에서 활발하게 사용되고 있었다.

“요즘에는 웹으로 표출하고 검색하고, (중략) 이것들은 전문 소프트웨어들이 있어요. ArcGIS도 많이 쓰고, QuantumGIS라고 해서 오픈소스 쪽도 많이 쓰고요. 커머셜한 것은 ArcGIS이고…”(P4, 40대, 경력 19년차)

컴퓨터공학 전공으로 인문사회 분야 데이터를 관리하는 피면담자의 경우, 멀티미디어와 이미지, 텍스트 데이터를 대상으로 데이터마이닝과 마인드마이닝 분석을 적용한다고 응답했다. 도출된 분석 결과를 웹서비스로 제공해 이용자 편의를 도모하려는 노력의 일환이다. 인문사회 분야의 피면담자는 데이터 분석 도구로 주로 SPSS 통계 패키지를 제공한다고 응답했다.

피면담자 대부분은 데이터 관리 시스템 구축 과정에서 데이터 표준화 업무를 수행한다고 응답했다. 표준화 활동은 데이터 수집·시스템 입력·관리 같은 보존 전략 및 절차를 정의하고, 이를 바탕으로 워크플로우를 설계하고 데이터 특성을 고려한 적절한 메타데이터를 선정하는 실무를 포함한다. 피면담자들은 데이터 생산자가 데이터를 수집하거나 생산해서 시스템에 입력하는 활동을 콘텐츠 생산 활동으로 정의했으며, 자신들은 시스템 구조에 대한 전반적인 이해를 바탕으로 콘텐츠 생산 과정에 필요한 규칙 및 준용할 메타데이터와 같은 자료를 데이터 생산자에게 제공한다고 응답했다.

“데이터 생산자가 하는 활동은 메타데이터를 만

드는 것이기 때문에 실제로 콘텐츠를 채우는 것이라고 할 수 있어요. 데이터 표준은 시스템 담당자인 제가 만드는 것이고요. 목록을 작성할 때도 제가 기본적으로는 목록 규칙을 만들어 드려요.”(P1, 40대, 경력 12년차)

기술에 필요한 메타데이터를 준용하는지 여부는 5곳의 기관에서 모두 준용한다고 답변했다. 준용하는 메타데이터 유형은 기관별로 모두 달랐다. 인문사회 분야에서는 ISAD(G), DDI(The Data Documentation Initiative), DC(Dublin Core), KCR4(한국목록규칙 4판)을 준용하고 있었다. 지질자원 관리 기관의 경우 지리정보시스템(Geographic Information System; 이하 GIS) 표준과 세부 주제 분야별 표준을 함께 준용할 계획이라고 응답했다. 그러나 메타데이터를 준용하는 정도에서는 기관별로 차이를 보였다. 피면담자 중 일부는 2종의 메타데이터 표준을 상호 보완하는 개념으로 함께 사용했다. 과학기술 분야의 경우는 GIS 데이터 표준을 기반으로 하고 추가적으로 환경, 광물, 해양 등 세부 주제 분야별 표준을 준용해 기술할 예정이라고 응답했다. 인문사회 분야 소속의 피면담자도 기록관리 메타데이터 표준과 도서관의 메타데이터 표준을 복합적으로 준용한다고 답변했으며, 그 목적은 데이터 기술을 더욱 상세하게 하고 이를 통해 검색 성공률을 높여 이용자 요구에 부응하고자 함이라고 했다.

“ISAD(G)를 기본 틀로 사용하고 있어요. 그런데 우리나라 이용자들 성향은 검색했을 때 원하는 정보가 바로 검색되는 것을 좋아해서 조금 더 상세한 메타데이터가 필요했어요. 그래서 도

서관 쪽 KCR4를 추가적으로 준용했어요.”(P1, 40대, 경력 12년차)

4.2.3 계획·관리·평가 활동

대다수의 피면담자는 서비스, 프로젝트, 프로그램을 계획하고 관리하고 평가하는 업무를 수행한다고 응답했다. 주로 기관 내부에서 데이터베이스로 구축한 데이터를 재사용하는 것에 목적을 두고, 새로운 서비스를 계획하고 개발하는 업무가 이루어지고 있었다. 서비스 유형으로는 콘텐츠 서비스 및 연구 지원 서비스가 있었으며, 이를 개발하는 업무를 수행했거나 계획 중이라고 응답했다. 콘텐츠 서비스의 경우 1차 생산된 원데이터를 구축하는 업무에서, 더 나아가 구축된 데이터를 활용한 결과물을 재수집해서 새로운 데이터 컬렉션을 구축한다는 답변도 있었다. 인문사회 분야 피면담자의 경우, 자관의 데이터를 활용해서 연구를 진행해 학술 및 학위논문을 출판하면 2차 생산물로 간주해 다시 기탁 받아 공개하는 서비스를 시행하고 있다고 응답했다.

“저희 쪽 자료를 활용해서 학위논문이나 학술논문을 작성하는 경우가 있어요. 그런 분들에게도 저작권 동의를 받고 기탁을 받으려고 하죠.”(P2, 30대, 경력 8년차)

새로운 서비스 개발에 앞서 기관 내부 및 외부 이해관계자들의 요구를 파악하는 작업이 수행되었다. 서비스 요구를 파악하기 위한 방법으로는 인식 조사 같은 설문 조사를 실시하거나, 시스템 이용자 평가 등이 있었다. 설문 조사의 경우, 데이터 생산자를 대상으로 데이터 공개

및 통합 관리에 대한 인식 조사와 함께 데이터 공개에 따른 적절한 보상 체계를 문의했다고 답했다. 설문 조사를 통해 데이터 공유를 촉진하는 보상 체계 및 유인책을 개발하려는 노력을 하고 있었다.

“저희가 설문 조사도 하고... 왜 데이터 관리가 안 되는지, 왜 데이터를 안 내놓으려하는지... (중략) 데이터는 생산자 자기 것이라는 인식이 있어서... (중략) 앱하고 설정도 해야 하고... 공개할 유인책을 마련하려고 하고 있고... 여러 가지 평가 제도라든지... 데이터 저널 같은 것도 만들려고...”(P4, 40대, 경력 19년차)

4.2.4 최신 동향 분석

피면담자들은 데이터 큐레이션 및 관리 실무에 대한 최신 기술 및 서비스, 연구 동향을 살펴보는 업무를 수행하고 있었다. 참조 방법은 주로 관련 기관 웹사이트 및 최신 논문을 살펴보는 것이나 관계 기관을 직접 방문하는 것이다. 관련 기관 웹사이트 및 최신 논문을 참조한다고 응답한 경우 피면담자 본인 스스로 정보를 수집하기도 했고, 관련 연구 용역에 참여해 최신 정보를 수집하고 동향을 파악한다는 답변도 있었다. 데이터 큐레이션 관련 최신 동향 파악에 주로 참조하는 자료원은, 해당 주제 분야의 선진적인 사례라고 여겨지는 국외 기관의 웹사이트를 주기적으로 확인한다고 응답했다. 인문사회 분야는 히스토리 센터, 박물관, 기록관의 웹사이트와 대표적인 데이터 아카이브의 웹사이트를 참조하고 있었다. 과학기술 분야 지질자원 기관의 실무자는 해외의 지질자원 데이터 관리 기관을 참조한다고 답변했다.

“(유관 주제 분야 사이트에) 본인들 큐레이션에 대해서 소개하는 페이지가 있어요. 아무래도 선진화된 기관이라고 여겨지니까 많이 참조를 하려고 노력하죠. 번역도 했고...”(P2, 30대, 경력 8년차)

동향 분석의 주 목적은 서비스 내용과 실무 사례 파악에 있었다. 웹사이트와 함께 관련 국내외 논문, 한국콘텐츠진흥원에서 발간하는 연구 동향 분석 보고서 등을 참조한다는 답변도 있었다. 메타데이터 준용과 같은 표준화 업무 및 실무 사례 참조가 필요한 경우 선행 연구자들의 논문을 참조한다고 응답했다. 이와 함께 서비스 구축 사업 시행 단계에서 참조할 대상 기관을 선정해 직접 방문하는 경우도 있었다. 국내 기관 및 국외 기관을 방문해서 데이터 관리 사례를 조사했다.

“여러 군데를 방문했어요. 미국 워싱턴대학교, 여기는 도서관에서 (아카이빙을) 해요. 그리고 자연사박물관, 벨기에의 국립아카이브센터... 그런 곳들을 다 방문했어요.”(P3-1, 60대, 경력 25년차)

4.2.5 서비스 제공

이용자 서비스와 관련해서 피면담자들은 교육과 훈련, 상담을 수행하고 있었다. 기관 내부에서는 업무가 내부 데이터베이스를 기반으로 수행되기 때문에 내부 인력에게 데이터베이스 이용 및 데이터 기술에 대한 교육을 제공할 필요가 있다고 응답했다. 이와 관련해서 피면담자 대부분이 이용 매뉴얼을 작성해서 제공하는 업무를 하고 있었다. 그러나 이용자 교육을 위해

정기적인 교육을 제공하고 있지는 않았다. 기관 외부 교육의 경우 데이터 생산자로부터 데이터를 기탁받을 때 데이터 보존에 필요한 교육을 지원한다고 응답했다. 또한 구축된 데이터를 재 사용하는 이용자에게는 활용 교육 및 상담을 제공했다. 피면담자 중 데이터 이용과 분석 서비스를 전문적으로 제공하는 기관 소속인 경우, 연구도서관에서 요청이 발생하면 데이터베이스 이용 교육과 데이터 서비스 관련 세미나, 특강 등을 진행한다고 응답했다. 이와 함께 관련 분야 교육 기관에 데이터 관리 교육의 필요성을 전달하고 교육 과정 개발을 제안했다고 응답한 피면담자도 있었다.

“데이터 관리에 대한 교육 같은 게 필요해서... 저희 쪽에 (관련) 교육센터가 있거든요. 그쪽과 연계를 해서 데이터 관리를 어떻게 할 것인지... 중요성과 필요성(을 알리고), 품질 관리 등은 어떻게 할 것인지... 전반적인 교육 프로그램을 만들 필요가 있겠다고 제안을 한 상태예요.”(P4, 40대, 경력 19년)

5. 논의

본 연구는 데이터 큐레이터 구인 공고와 실무자 면담을 통해서 데이터 큐레이터의 핵심 직무 요건을 데이터 관리 실무의 측면에서 규명하고자 했다. 국외 데이터 큐레이션 관련 구인 공고와 국내 데이터 관리 실무자 심층 면담을 기반으로 내용 분석을 실시해 데이터 큐레이터 직무의 특성과 핵심 직무 요건을 분석했다. 이들 분석에서 도출된 특성을 커뮤니케이션, 정보기술

활용, 데이터 관리 지식, 서비스 제공의 4가지 측면을 살펴보면 다음과 같은 특성을 논의할 수 있다.

첫째, 데이터 큐레이터 직무는 데이터 관리와 관련된 다양한 이해관계자들과의 활발한 커뮤니케이션 활동이 요구되고 있었다. 국외 구인 공고 분석의 경우 의사소통 및 대인관계 요건에서 원활한 커뮤니케이션 능력이 67%로 가장 선호되었다. 또한 국내 데이터 관리 실무자들 역시 데이터 생산자 및 정보기술 개발자 등 다양한 이해관계자와 협업하고 있으며, 이해관계자가 속한 주제 지식 및 직무 특성의 차이들을 이해하고 이에 기반을 두어 의사소통 하는 것을 중요하게 여기는 것으로 나타났다. 특히 피면담자들은 커뮤니케이션을 통해 이해관계자를 설득하고 데이터 관리에 대한 공감대를 형성하는 것이 가장 어렵지만 중요한 업무라고 응답했다. 이러한 활동은 상이한 주제 분야와 데이터 관리 인식을 가진 데이터 생산자와 정보기술 개발자와의 사이에서 효율적으로 데이터 관리 시스템을 구축하고 운영하는 것을 포함한다. 본 연구에서 밝힌 커뮤니케이션의 중요성은 데이터 생산을 위해서는 학문과 정보기술의 접목이 요구되며, 접목에 있어 데이터 큐레이터가 두 영역을 잇는 교량 역할을 해야 한다고 주장한 Stanton et al.(2011)의 연구와 맥락을 같이한다고 볼 수 있다.

둘째, 데이터 큐레이터 직무는 정보기술에 대한 지식을 바탕으로 데이터 관리 시스템을 구축하고 운영하며, 변화하는 기술 환경을 이해하고 최신성을 유지하는 역할이 요구된다. 국외 데이터 큐레이터 관련 구인 공고 분석 결과 7가지 직무 요건 중에서 큐레이션 관련 기술 활용 요

건이 가장 높게 요구되고 있었다. 또한 국내 데이터 큐레이션 실무자들도 기관 내부 데이터 관리 시스템을 구축 하고 운영하는 실무를 수행하고 있어서, 데이터 큐레이터에게 있어 정보기술에 대한 지식이 업무 수행에 중요하게 요구되고 있음을 알 수 있었다. 특히 국외 구인 공고와 국내 실무자 면담 분석에서 데이터 관리 시스템 구축에는 데이터베이스 관리 시스템 지식과 프로그래밍 언어, 마크업 언어 지식과 같은 정보기술에 대한 이해를 공통적으로 요구하는 것으로 나타났다. 해당 언어를 활용해 실제로 시스템을 코딩하는 것은 정보기술 개발자이지만, 데이터 큐레이터에게도 시스템 개발에 대한 청사진과 방향성을 제공할 수 있는 정도의 지식이 필수적으로 요구되었다. 그러나 실무가 생산자 및 개발자 중에서 어느 분야와 더 밀접한 관계를 맺고 있느냐에 따라 주제 분야 및 연구 과정에 대한 이해도가 더 중요한지, 정보기술 관련 지식이 더 중요하게 요구되는지에 대한 차이를 보였다. “그래서 아키비스트팀과 소프트웨어 개발팀은 항상 같이 가야 해요.”(P3-1, 60대, 경력 25년차)라고 응답한 피면담자의 의견은 데이터 큐레이션에 있어 큐레이션 전략 및 지식과 큐레이션 관련 정보기술을 동시에 한 사람이 모두 보유하는 것이 쉽지 않음을 나타내준다. 중장기적으로 기관 내 데이터 큐레이션을 지속하기 위해서는 정보기술 개발자와 데이터 관리자가 항상 협업을 수행하는 것이 중요하다는 점을 강조하고 있다. 이와 함께 빠르게 변화하는 정보기술 환경 속에서 큐레이션 이슈 및 기술 변화에 대한 최신 동향을 수집하고 분석하는 요건이 국외 사례 및 국내 실무자 심층 면담에서도 중요하게 요구되는 것으로 나타났다. 국내

데이터 큐레이션 실무자들은 최신 데이터 큐레이션 서비스 및 선진 실무 사례를 파악하는 업무를 수행했다. 주로 해당 분야의 대표적인 기관 웹사이트를 주기적으로 방문하거나 관련 연구 용역에 참여하고 최신 논문을 참고했다. 기관 내부의 지원이 있을 경우 선진 실무가 수행되는 국내외 기관을 방문하기도 했다.

셋째, 데이터 큐레이터는 큐레이션 및 보존 활동에 대한 지식과 전략이 요구되는 것으로 나타났다. 국외 구인 공고 분석에서는 7가지 직무 요건 중에서 큐레이션 및 보존 활동 요건이 세 번째로 높게 나타났다. 큐레이션 및 보존 활동 요건의 세부 항목을 살펴보면 큐레이션 전략 및 절차에 대한 지식과 메타데이터 표준 준용이 전체의 63%를 차지해 데이터 관리에 있어 큐레이션 전략을 이해하고 관련 메타데이터를 준용하는 활동이 핵심적으로 요구됨을 알 수 있다. 특히 메타데이터 표준 지식은 4년 동안 매년 20%를 차지해 그 중요도를 파악할 수 있다. 국내 데이터 관리 실무자들도 표준화된 데이터 관리 인프라스트럭처를 구축하기 위해 업무 절차를 정의하고 해당 주제 분야의 표준 메타데이터를 준용하고 있었다. 메타데이터 표준의 경우, 이용자의 만족도와 상호운용성을 높이기 위해 기록관리 및 도서관 표준 메타데이터를 추가적으로 적용하여 상세하게 기술하고 있었다. 또한 다양한 데이터 유형 및 특징, 생산되는 맥락에 대한 지식이 중요하게 요구되는 것으로 나타났다. 그러나 데이터 관리에 대한 전략에 대해서는 국외와 국내의 요구에 있어 차이를 보였다. 국외의 경우 큐레이션 및 보존 활동 요건의 세부 항목 중 큐레이션 전략 및 절차 지식이 43%(290건)로 4년 합계 기준으로 가장 높은

빈도를 보여 데이터 관리 전략에 대한 이해와 지식을 중요하게 요구하고 있었다. 반면 국내의 경우는 구축한 데이터의 재사용에 대한 전략 정도로 이해하거나 데이터 큐레이션 전략은 이론적인 내용이라고 답변해, 전문적인 데이터 관리 지식과 전략에 대한 요구가 높지 않은 것으로 나타났다. “(데이터 큐레이션) 절차가 굉장히 이론적인 얘기인 거죠. 실무를 하다 보면 할 수밖에 없는 얘기거든요... 업무 자체가 그것인 거죠.”(P2, 30대, 경력 8년차)라고 응답한 피면담자의 의견은 데이터 큐레이션에 관한 전략이 데이터 관리 업무에 미치는 효용에 대한 관심이 낮은 것으로 이해할 수 있다.

넷째, 데이터 큐레이터 직무는 원활한 데이터 관리와 그것의 재사용을 촉진하기 위해 다양한 서비스를 계획하고 제공하는 활동이 요구되고 있었다. 국외 구인 공고 분석에서 서비스 요건 중 리에중 및 교육 등의 이용자 서비스를 제공하는 업무 수행이 요구되었다. 국내 데이터 관리 실무자들 또한 시스템 이용자들의 정보 요구를 파악하고 그에 맞는 서비스를 개발하는 업무를 수행하는 것으로 나타났다. 설문 조사를 실시해 데이터 생산자들의 데이터 공유 및 관리에 대한 인식을 조사하고 시스템의 기능을 개선하기 위해 이용 평가를 실시하고 있었다. 이와 더불어 데이터 공유를 촉진할 보상 체계 및 유인책 개발, 데이터 시각화 도구를 개발하는 등 다양한 서비스를 제공하려는 노력을 기울이고 있었다. 웹서비스도 소장 데이터를 제공하는 것뿐만 아니라 소장 데이터를 분석한 통계 자료와 시각화 자료를 함께 제공해서 데이터 활용을 촉진하려는 노력을 하고 있었다. 이와 함께 데이터 관리 및 이용에 관한 교육 활동도

수행하는 것으로 나타났다. 국외 구인 공고 분석에서 이용자 지원 서비스의 비중은 감소하는 반면 트레이닝 및 컨설팅에 대한 요구가 증가해, 데이터 관리에 대한 교육 서비스의 중요성을 확인할 수 있다. 국내의 경우 현재는 보존된 데이터를 활용해 연구를 수행하려는 이용자에 한해서 교육이 제공되고 있었으며, 제공하는 기관도 데이터 서비스를 전문적으로 수행하는 기관의 경우로 한정되어 있었다. 그러나 국내 데이터 관리 실무자들의 경우 “온라인으로 생산자들이 데이터를 입력할 때, 규정된 형식에 맞지 않게 데이터를 입력하는 게 업무를 어렵게 만듭니다.”(P3-2, 30대, 경력 8년차)와 같은 답변을 통해, 데이터 관리 지식의 부족으로 규정된 형식에 맞지 않게 데이터를 입력해서 실무자가 일일이 수정해 줘야 하는 일이 발생하기 때문에 이를 방지하기 위해 데이터 관리 교육이 필요함을 인식하고 있었다.

6. 결론

빠른 속도로 발전하는 정보기술과 인터넷 환경으로 인해 데이터는 다양한 학문과 사회 현상의 중심에 놓여있다. 데이터를 효율적으로 관리하고 보존하여 재사용이 가능하도록 하는 실무를 담당하는 데이터 큐레이터에 대한 관심 역시 증가하고 있으나, 아직까지는 데이터 큐레이터의 개념 및 직무에 대하여 명확하게 규정되지 않은 실정이다. 본 연구는 국내외에서 실무를 살펴볼 수 있는 구인 공고와 실무자 면담을 통해, 데이터 큐레이터 직무를 규명하고자 하였다.

이러한 분석결과를 통해 본 연구는 문헌정보학 교과과정에 있어서 전문적인 대규모 데이터 관리에 대한 데이터 큐레이션 교육과정 설계를 제안할 수 있다. 현재 국내 문헌정보학 교과 과정 중에서 대규모 데이터 관리 및 큐레이션과 관련해 전문적인 교육을 제공하고 있는 곳은 거의 없는 실정이다(이혜영 2014). 또한 면담에 참여했던 실무자들도 데이터 관리에 대한 전략 및 지식을 습득해 본 경험이 없다고 응답했다. 반면 국외 구인 공고 분석에서는 전문적인 데이터 큐레이션 및 보존 전략에 대한 지식을 높게 요구하고 있었다. 또한 주제 분야 별로 데이터 큐레이터의 직무를 담당하는 실무자의 전공과 정보기술 활용에 대한 요구 정도가 상이한 실정이었다. 그렇지만 앞서 분석한 결과와 같이 이해관계자와 공감대를 형성하고 데이터 관리에 대한 전략과 정보기술 및 동향에 대해 최신성을 유지하며, 이용자의 요구를 분석하고 이에 맞는 서비스를 개발하는 업무들은 데이터 관리 업무가 단순히 정보기술의 활용 여부만으로 규정될 수 있는 직무라고 판단하기는 어렵다. 데이터 큐레이터 직무가 갖는 다양한 업무 활동들은 정보기술과 함께 데이터를 관리하는 전략에 기반하고 있기 때문이다. 또한 주제 분야별로 적용되는 메타데이터 표준과 데이터 생애 주기 전략 활용과 같은 다양한 측면들이 존재한다. 따라서 위와 같은 현황들을 종합하면 국내에도 대규모 디지털 데이터 큐레이션에 대한 통합적이고 전문적인 교육의 제공이 필요하다고 사료된다. 국

외의 경우 일부 문헌정보학 대학원에서 데이터 큐레이션 교과 과정 프레임워크를 설계해서 교육에 반영하고자 노력하고 있으며, 전공 분야가 문헌정보학이 아니더라도 다양한 주제 분야의 데이터 관리 실무자들을 추가적으로 교육할 수 있는 전문 프로그램을 제공하고 있다. 또한 문화유산 영역의 데이터 관리 기관에서도 데이터 큐레이션을 직무 훈련 형태로도 제공하고 있어, 데이터 큐레이션 교육은 학문 영역과 실무 영역에서 다양한 관점으로 제공되고 있다. 이러한 전문적 교육의 제공은 직무의 전문성을 도모하는 동시에, 기관 내 데이터 관리에 대한 인식을 개선하고, 데이터 관리에 대한 전략을 세워서 효율적인 데이터 관리와 공유 활동을 촉진한다. 국내에서도 면담을 통해 일부 데이터 관리 기관에서 해당 분야 데이터 관리 교육을 제안한 사실을 파악했으나 현재까지 시행이 되고 있지는 않으며, 또한 특정 분야에만 국한 되어 있는 교육이기에 통합적인 데이터 관리 교육에 대한 요구를 충족시킬 수는 없다고 판단된다.

데이터의 규모가 증가하고, 대규모 데이터를 통해 의미를 도출하려는 최근의 연구 경향은 체계적인 데이터 관리에 대한 중요성을 환기시키고 있다. 이에 체계적인 데이터 관리를 수행하는 직무인 데이터 큐레이터의 직무 현황을 고찰해보고, 이를 통해 데이터에 대한 통합적인 관리 교육을 제안하는 것은 데이터 관리에 대한 인식을 고취시키고 실무에 대한 전문성을 높이는 데 일조할 수 있을 것으로 기대한다.

참 고 문 헌

- 곽승진, 최재황, 배경재, 정영미. 2011. 과학기술분야 디지털 콘텐츠의 아카이빙 정책 연구. 『정보관리 연구』, 42(4): 115-136.
- 김선태, 한선화, 이태영, 김용. 2010. 과학데이터 보존 및 활용모델에 관한 연구. 『한국비블리아학회지』, 21(4): 81-93.
- 김지현. 2014. 대학도서관의 연구데이터관리서비스에 관한 연구. 『한국비블리아학회지』, 25(3): 165-189.
- 신영란. 2012. 『인문사회 연구데이터 아카이브의 발전 방안에 관한 연구』. 석사학위논문. 이화여자대학교 정책과학대학원.
- 안영희, 박옥화. 2009. 대학도서관 서비스의 디지털 큐레이션 전략. 『한국도서관·정보학회지』, 40(4): 311-326.
- 안영희. 2011. 『디지털 큐레이션 서비스의 프레임워크 개발에 관한 연구』. 박사학위논문. 충남대학교 대학원, 문헌정보학과.
- 이혜영. 2014. 외국대학의 디지털 큐레이션 교육 프로그램 동향. 한국도서관·정보학회 동계학술발표회, 67-85.
- 한나은. 2014. 『대학도서관의 디지털 큐레이션 프로세스에 관한 연구』. 석사학위논문. 중앙대학교 대학원, 문헌정보학과.
- Choi, Y. and E. Rasmussen, 2009. "What Qualifications and Skills are Important for Digital Librarian Positions in Academic Libraries? A Job Advertisement Analysis." *The Journal of Academic Librarianship*, 35(5): 457-467.
- Corrall, S. 2012. "A Competency Framework for Digital Curation and Data Science." In: *Aligning Digital Preservation across Nations*, Amsterdam: Netherlands.
- Creamer, A., M. E. Morales, J. Crespo, D. Kafel, and E. R. Martin. 2012. "An Assessment of Needed Competencies to Promote the Data Curation and Management Librarianship of Health Sciences and Science and Technology Librarians in New England." *Journal of eScience Librarianship*, 1(1): 4.
- Dekker, R. 2006. "The Importance of Having Dataset." In *27th ATUL Conference, 22-25 May 2006*. IATUL.
- Gold, A. 2010. "Data Curation and Libraries: Short-term Developments, Long-term Prospects." *Office of the Dean (Library)*, 27.
- Heidorn, P. B. 2011. "The Emerging Role of Libraries in Data Curation and e-science." *Journal of Library Administration*, 51(7-8): 662-672.

- Kim, J., E. Warga, and W. Moen. 2013. "Competencies Required for Digital Curation: An Analysis of Job Advertisements." *International Journal of Digital Curation*, 8(1): 66-83.
- Kruse, F. and J. B. Thestrup. 2014. "Research Libraries' New Role in Research Data Management, Current Trends and Visions in Denmark." *Liber Quarterly*, 23(4): 310-335.
- Lord, P., A. MacDonald, L. Lyon, and D. Giarretta. 2004. "From Data Deluge to Data Curation." In *Proceedings of the UK e-science All Hands meeting*, 371-357.
- Martinez-Uribe, L. M. and S. MacDonald. 2008. "A New Role for Academic Librarians: Data Curation." *Profesional de la Informacion*, 17: 273-280.
- Parkinson, H. 2008. "A Data Scientist's Perspective on Roles and Responsibilities in Data Curation." *Research Data Management Forum 2008*. [Presentation slides].
 <<http://www.dcc.ac.uk/sites/default/files/documents/RDMF/RDMF2/04%20Parkinson.pdf>>.
- Stanton, J. M., Y. Kim, M. Oakleaf, R. D. Lankes, P. Gandel, D. Cogburn, and E. D. Liddy. 2011. "Education for eScience Professionals: Job Analysis, Curriculum Guidance, and Program Considerations." *Journal of Education for Library and Information Science*, 52(2): 79-94.
- Tamaro, A. M., S. Ross, and V. Casarosa. 2014. "Research Data Curator: The Competencies Gap." *BOBCATSSS 2014 Proceedings*, 1(1): 95-100.

[웹사이트]

- 북미연구도서관협회(Association of Research Libraries) [cited 2015.3.12]. <<http://www.arl.org/>>.
- ASIS&T Jobline [cited 2015.3.12]. <<http://asis-jobs.jobtarget.com/jobseeker/search/results/>>.
- Code4lib [cited 2015.3.16]. <<http://jobs.code4lib.org/>>.
- Digital Curation Centre [cited 2015.3.1]. <<http://www.dcc.ac.uk/>>.
- Digital Curation Exchange [cited 2015.3.12]. <<http://digitalcurationexchange.org/>>.

• 국문 참고자료의 영어 표기

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

- Ahn, Y. 2011. *A Study of the Framework Development of Digital Curation Service*. Ph. D. Dissertation, Department of Library and Information Science, Graduate School, Chungnam National University.
- Ahn, Y. and O. Park. 2009. "Digital Curation Strategy for University Library Service." *Journal of Korean Library and Information Science Society*, 40(4): 311-326.
- Han, N. 2014. *A Study on Digital Curation Process in Academic Libraries*. Master's Thesis,

Department of Library and Information Science, The Graduate School of Chung-Ang University.

- Kim, J. 2014. "A Study on Research Data Management Services of Research University Libraries in the U. S." *Journal of the Korean Biblia Society for Library and Information Science*, 25(3): 165-189.
- Kim, S., S. Hahn, T. Lee, and Y. Kim. 2010. "A Study on a Model for Using and Preserving Scientific Data." *Journal of the Korean Biblia Society for Library and Information Science*, 21(4): 81-93.
- Kwak, S., C. Choi, K. Bae, and Y. Jung. 2011. "A Study on the Policy on Digital Contents Archiving in the Field of Science and Technology." *Journal of Information Management*, 42(4): 115-136.
- Rhee, H. 2014. "Digital Curation Education Program Trends of Foreign Universities." *Korean Library and Information Science Society Winter Conference*, 2014(11): 67-85.
- Shin, Y. 2012. *A Study on the Development Plan for Humanities and Social Sciences Research Data Archive. Master's Thesis*, The Graduate School of Policy Science, Ewha Womans University.