

북미 대학도서관 연구데이터 관리 교육 프로그램 내용 분석: 데이터 리터러시 세부 역량을 중심으로

A Content Analysis of Research Data Management Training Programs at the University Libraries in North America: Focusing on Data Literacy Competencies

김지현 (Jihyun Kim)*

초 록

본 연구에서는 북미에서 연구데이터 관리 서비스를 제공하는 121개 대학도서관 중 연구데이터 관리 교육 프로그램을 제공하는 51개 기관을 대상으로 제공되는 교육 프로그램의 내용을 12개 데이터 리터러시 세부 역량에 기반을 두어 분석하고 시사점을 제시하는 것을 목적으로 하였다. 내용 분석을 위해 집합 교육 프로그램의 제목 317개와 온라인 튜토리얼의 상위 목차 제목 42개를 수집하였으며 선행연구에서 제시된 12개 데이터 리터러시 세부 역량에 따라 코딩을 수행하였다. 집합 교육 프로그램 중에서는 데이터 처리 및 분석 역량에 대한 교육 프로그램이 가장 많은 것으로 나타났으며, 가장 많은 수의 기관에서 데이터 관리 및 조직 역량에 대한 교육을 제공하고 있었다. 데이터 시각화 및 표현은 집합 교육 프로그램 중에서 세 번째로 많이 다루어지는 역량이었다. 그러나 나머지 9개 역량에 대한 교육 프로그램은 매우 적은 것으로 나타나 교육 프로그램 내용이 특정 역량에 집중되어 있음을 알 수 있다. 집합 교육 없이 자체 개발한 온라인 튜토리얼을 제공하는 기관은 5곳이었으며 목차 제목을 분석한 결과 데이터 보존, 윤리 및 데이터 인용, 데이터 관리 및 조직 역량에 대한 교육 내용을 중점적으로 다루고 있어 집합 교육 프로그램에서 강조되는 역량과 차이를 보였다. 효과적인 연구데이터 관리 교육 프로그램 운영을 위해서는 대학도서관 사서들이 전통적으로 교육하고 강조해왔던 역량뿐만 아니라 데이터 처리와 분석, 데이터 시각화와 표현 등 연구자들의 연구 결과 도출에 필요한 데이터 리터러시 세부 역량에 대한 이해와 지원이 요청된다. 또한 연구데이터 관리 서비스와 관련된 사서들의 계속 교육을 지원하는 교육 자원의 개발도 필요할 것이다.

ABSTRACT

This study aimed to analyze the content of Records Data Management (RDM) training programs provided by 51 out of 121 university libraries in North America that implemented RDM services, and to provide implications from the results. For the content analysis, 317 titles of classroom training programs and 42 headings at the highest level from the tables of content of online tutorials were collected and coded based on 12 data literacy competencies identified from previous studies. Among classroom training programs, those regarding data processing and analysis competency were offered the most. The highest number of the libraries provided classroom training programs in relation to data management and organization competency. The third most classroom training programs dealt with data visualization and representation competency. However, each of the remaining 9 competencies was covered by only a few classroom training programs, and this implied that classroom training programs focused on the particular data literacy competencies. There were five university libraries that developed and provided their own online tutorials. The analysis of the headings showed that the competencies of data preservation, ethics and data citation, and data management and organization were mainly covered and the difference existed in the competencies stressed by the classroom training programs. For effective RDM training program, it is necessary to understand and support the education of data literacy competencies that researchers need to draw research results, in addition to competencies that university librarians traditionally have taught and emphasized. It is also needed to develop educational resources that support continuing education for the librarians involved in RDM services.

키워드: 연구데이터 관리 서비스, 교육 프로그램, 데이터 리터러시 세부 역량
research data management services, training programs, data literacy competencies

* 이화여자대학교 사회과학대학 문헌정보학과 부교수(kim.jh@ewha.ac.kr)

■ 논문접수일자: 2018년 10월 30일 ■ 최초심사일자: 2018년 11월 20일 ■ 게재확정일자: 2018년 12월 4일
■ 정보관리학회지, 35(4), 7-36, 2018. [http://dx.doi.org/10.3743/KOSIM.2018.35.4.007]

1. 서론

대학 구성원들의 연구 활동을 지원하는 대학 도서관의 전통적인 역할은 여전히 유효하고 가치 있는 활동이며 이러한 역할을 강화하고 확장하는 차원에서 시작된 연구데이터 관리 서비스는 지난 10년 간 미국, 호주, 영국 등 다양한 해외의 대학도서관에서 지속적으로 논의·개발되어 왔다(Brown, Wolski, & Richardson, 2015; Faniel & Connaway, 2018; Tenopir, Talja, Horstmann, Late, Hughes, Pollock, Schmidt, Baird, Sandusky, & Allard, 2017). 오픈 사이언스, 연구비지원기관의 데이터 제출 의무 정책, 재현 가능한 연구 및 윤리적인 연구 수행이 강조되고 있는 현실에서 연구데이터 관리 서비스의 필요성은 더욱 높아지고 있는 실정이다.

연구데이터 관리 서비스의 영역은 크게 세 가지로 교육, 큐레이션 및 전문성(expertise)으로 요약할 수 있다. 즉, 교육을 통해 연구데이터 관리의 중요성에 대한 인식을 높이고 관련 지식과 능력을 배양하며 이와 관련된 도구 및 자원으로 안내하는 역할을 수행할 수 있다. 큐레이션은 연구 주기를 통해 이루어지는 데이터 관리 활동을 지원하는 기술적 인프라 구축 및 관련 서비스 제공을 의미한다. 전문성 영역은 연구데이터 관리에 있어 발생할 수 있는 특정한 문제점들에 대해 맞춤형 해결 방안을 제시하고 의사 결정을 지원할 수 있는 서비스를 포함하고 있다(Bryant, Lavoie, & Malpas, 2017).

본 연구에서는 이러한 연구데이터 관리 서비스의 영역 중 교육의 영역에 초점을 맞추어 대학도서관의 중요한 기능 중 하나인 정보 리터러시 교육에서 확장된 것이라고 볼 수 있는 데

이터 리터러시 교육의 현황을 조사하고자 하였다. 이를 위해 북미 대학도서관에서 운영하는 연구데이터 관리 교육 프로그램의 내용을 집합 교육과 온라인 교육의 두 가지 제공 방식에 따라 구분하여 살펴보았으며 선행연구에서 논의된 데이터 리터러시 세부 역량을 중심으로 교육 내용을 분석하였다. 본 연구에서 데이터 리터러시는 연구데이터를 생성, 수집, 기술, 관리, 평가, 보존, 이용 및 재이용하는 능력이며 연구자 또는 데이터를 다루는 실무자의 능력을 의미한다. 데이터 리터러시 함양은 연구데이터 관리 교육 프로그램의 궁극적인 목적이므로 해당 교육 프로그램에서 어떠한 데이터 리터러시 세부 역량이 다루어지고 있는지를 논의하였으며 구체적인 연구의 목적 및 방법과 범위는 다음과 같다.

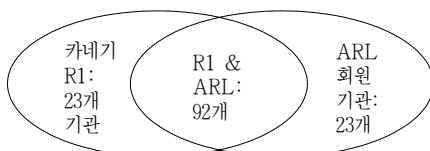
1.1 연구의 목적

본 연구에서는 북미 대학도서관의 연구데이터 관리 교육 프로그램에서 다루고 있는 데이터 리터러시 세부 역량을 프로그램의 제목 또는 목차 제목에 기반을 두어 분석하고 시사점을 제시하는 것을 목적으로 하였다. 본 연구에서 의미하는 교육 프로그램은 집합 교육 형태로 제공되는 워크숍이나 세미나, 혹은 기관에서 자체적으로 개발한 온라인 튜토리얼을 포함한다. 교육 제공 방식에 있어서는 집합 교육과 온라인 교육의 두 가지 방식으로 구분하여 살펴보았다. 이를 바탕으로 북미의 연구데이터 관리 교육 프로그램에서 빈번하게 다루어지는 데이터 리터러시 세부 역량과 그렇지 않은 역량은 무엇인지를 연구 문제로 제시하고 분석을 수행하였다.

1.2 연구의 범위와 방법

조사대상 기관을 선정하기 위해 먼저 본 연구에서는 미국 고등교육기관 분류체계인 카네기 분류(Carnegie Classification of Institutions of Higher Education)의 기본 분류(Basic Classification)에서 박사학위를 수여하면서 가장 활발히 연구 활동을 수행하는 R1(Doctorate Universities: Higher Research Activities)에 속하는 대학 115곳의 도서관을 조사대상 기관 범위에 포함하였다. 또한 연구도서관협회(Association of Research Libraries)의 회원도서관 124개 중 북미의 대학도서관 115개(미국 101개, 캐나다 14개)를 함께 포함하여 조사하였다. 카네기 R1 대학과 ARL 회원도서관의 모기관 모두 연구 활동이 왕성하게 이루어지는 대학들이므로 연구데이터 관리 서비스 및 교육 프로그램이 운영되고 있을 가능성이 높다고 판단하였다.

이 중 92개 대학이 카네기 R1 분류에 속하면서 ARL 회원도서관이 속한 기관임을 확인하였다. 23개 대학도서관은 모기관이 카네기 R1에 속하지만 ARL 회원도서관은 아니었으며, 또 다른 23개 도서관은 ARL 회원도서관이지만 모기관이 카네기 R1 대학은 아닌 것으로 나타났다. 따라서 본 연구의 조사대상기관은 <그림 1>과 같이 총 138개 대학도서관(미국 124개, 캐나다 14개)을 포함하였다.



<그림 1> 조사대상 기관의 카네기 R1 및 ARL 회원 여부

우선 조사대상 대학도서관의 연구데이터 관리 서비스 제공 여부를 확인하기 위해 연구데이터 관리 서비스를 설명하는 웹사이트 혹은 도서관가이드가 존재하는지, 그리고 해당서비스의 담당자 혹은 담당 팀이 존재하는지 여부를 조사하였다. 그 결과 138개 대학도서관 중 121개 대학도서관이 앞서 언급한 두 가지 조건을 충족하면서 연구데이터 관리 서비스를 제공하는 것으로 확인되었다. 이들 121개 기관 중 연구데이터 관리 서비스 웹사이트 혹은 도서관가이드에서 교육 프로그램에 관한 메뉴를 제공하는 기관은 모두 51개 기관인 것으로 나타났다. 교육 프로그램 메뉴 제목은 education, workshops, training, tutorials, events 등 교육이나 행사 관련 용어로 제시되어 있었다.

이들 51개 기관 중 전통적인 교육 방식인 집합교육의 형태로 교육 프로그램을 제공하는 기관은 43개로 다수를 차지하였고, 집합 교육 없이 자체 개발한 온라인 튜토리얼만을 제공하는 기관은 5개 기관, 집합 교육 없이 외부 기관의 온라인 튜토리얼 링크를 제공하는 기관은 3곳이었다. 집합교육과 외부 기관의 온라인 튜토리얼 링크를 함께 제공하는 기관도 43개 기관 중 9곳인 것으로 나타났다([부록 1] 참조).

본 연구에서는 집합 교육과 온라인 교육 두 가지 교육 제공 방식에 따라 연구데이터 관리 교육 프로그램을 구분하고 각각의 방식에서 제공되는 교육 프로그램의 제목을 수집하여 선행 연구에서 제시된 12개 데이터 리터러시 세부 역량으로 코딩을 실시하였다. 이들 역량은 1) 데이터 검색 및 수집; 2) 데이터베이스 및 데이터 포맷; 3) 데이터 변환 및 상호운용성; 4) 데이터 처리 및 분석; 5) 데이터 시각화 및 표

현; 6) 데이터 관리 및 조직; 7) 데이터 품질 및 도큐멘테이션; 8) 메타데이터와 데이터 기술; 9) 데이터 관리 및 공유 문화; 10) 윤리 및 데이터 인용; 11) 데이터 큐레이션; 12) 데이터 보존을 포함한다(Carlson, Fosmire, Miller, & Nelson, 2011). 집합 교육의 경우 개별 워크숍 또는 세미나의 제목 317개에 대한 코딩을 실시하였고 자체 개발된 온라인 튜토리얼의 경우 목차에서 제시된 가장 상위의 제목 42개를 중심으로 코딩을 수행하였다. 이후 각각의 데이터 리터러시 세부 역량에 관련된 교육 프로그램들을 세부적인 교육 내용에 따라 추가적으로 범주화하여 분석을 실시하였다.

2. 데이터 리터러시의 개념 및 역량

2.1 데이터 리터러시: 정의

Carlson 외(2011)는 데이터 정보 리터러시(data information literacy)를 이전부터 존재하였던 개념인 데이터 리터러시(data literacy), 통계 리터러시(statistical literacy) 및 정보 리터러시(information literacy)와의 연관성을 중심으로 논의하였다. 저자들은 해당 개념들을 통계 교육 관련 논문에서 인용하였으므로 데이터 리터러시와 통계 리터러시는 모두 통계와 연관된 개념으로 정의되고 있다. 즉 데이터 리터러시는 그래프와 차트를 적절히 읽고 데이터로부터 올바른 결론을 도출하며 데이터가 부적절하게 이용되었을 때 이를 인지하는 능력을 의미한다. 통계 리터러시는 일상적인 매체, 그래프, 표, 설문,

연구에서 제시된 요약된 통계를 읽고 해석하는 능력을 나타낸다. 정보 리터러시는 정보를 읽고 해석하고 평가하면서 어떠한 개념, 주장 및 논쟁에 대해 비판적으로 생각하는 능력을 의미한다. 이러한 세 가지 개념은 다시 비판적 사고 능력을 구성하는 계층적 요소들로 정의되었는데 데이터 리터러시는 통계 리터러시의 전제조건으로, 그리고 통계 리터러시는 정보 리터러시의 필수조건으로 제시되었다.

데이터 정보 리터러시와 유사한 개념으로 Qin와 D'Ignazio(2010)는 과학데이터 리터러시(science data literacy)의 개념을 논의하였다. 이는 데이터를 수집, 처리, 관리, 평가하고 과학적 탐구를 위해 데이터를 이용하는 능력을 의미한다. 해당 저자들은 과학데이터 리터러시가 과학 연구 종사자들뿐만 아니라 과학데이터 관리 전문가들에게도 필요한 능력임을 언급하였다. 또한 Schneider(2013)는 연구데이터 리터러시(research data literacy)라는 용어를 사용하였는데 이는 정보 리터러시의 하위 개념으로 연구데이터를 생성, 관리, 재이용하는 능력을 의미한다. 그는 정보 리터러시 핵심 능력을 바탕으로 연구데이터 리터러시와 관련된 8가지 능력으로 연구데이터의 식별, 조사(scope), 계획, 저장, 보호, 평가, 관리, 제공을 제안하였다.

한편 Shorish(2015)는 과학데이터 리터러시와 같은 용어에서도 나타나듯이 데이터 정보 리터러시에 대한 논의가 과학 분야에만 한정되어 있음을 지적하면서 인문학 및 사회과학 분야에도 적용될 수 있음을 제안하였다. 한상우(2017) 역시 디지털 인문학의 영향으로 다양한 형태의 데이터가 생산, 활용되고 있는 인문학 분야의 데이터 리터러시 모형 연구가 필요함을 주장하였다.

이와 더불어 Shorish는 정보 리터러시가 데이터 정보 리터러시와 밀접히 연관되어 있으며 대학도서관 사서들이 데이터 정보 리터러시 교육에 적극적으로 관여할 것을 제안하였다. 그는 미국대학도서관협회(Association of College and Research Libraries: ACRL)에서 2013년 제안한 '고등교육에서의 정보 리터러시를 위한 프레임워크(framework for information literacy for higher education)'가 데이터 생산자 및 이용자의 활동을 이해하는 문지방 개념(threshold concept)으로 활용될 수 있음을 제시하였다. 이는 기존의 정보 리터러시 개념이 데이터 리터러시를 이해하는 토대가 된다는 것을 나타낸다.

이재운(2015)은 문헌정보학 관점에서 중요한 데이터 리터러시의 개념으로 연구데이터 리터러시, 통계데이터 리터러시, 실무데이터 리터러시를 제시하였다. 연구데이터 리터러시와 통계데이터 리터러시는 선행연구에서 논의된 개념이며, 실무데이터 리터러시는 다양한 분야의 실무에서 생성된 데이터를 의사 결정이나 실무 개선을 위해 분석, 활용하는 능력을 의미한다.

이와 같이 다양한 측면에서 정의된 개념들을 바탕으로 Carlson 외(2011)는 기존 개념들에서 연구자의 정보 이용 및 분석 능력은 강조된 반면 정보 생산 능력은 간과되었음을 지적하였다. 따라서 데이터 정보 리터러시의 개념은 데이터 생산자, 그리고 데이터 이용자로서의 연구자들의 관점을 포괄해야 한다고 보았다. 이를 바탕으로 데이터 정보 리터러시를 기존의 리터러시 개념들을 기반으로 재통합된, 새롭게 등장하는 능력으로 정의하였다.

Koltay(2016)는 데이터 정보 리터러시, 과학 데이터 리터러시, 연구데이터 리터러시 등 여러

가지 용어로 제시되고 있는 데이터 리터러시를 통일된 용어로 정의할 것을 제안하였다. Kahn, Kim과 Chang(2017)은 선행연구에서 제시된 데이터 리터러시의 개념을 1) 데이터기반 의사결정(Data-Driven Decision-making)에 필요한 능력; 2) 연구데이터 관리 서비스를 지원하는 활동; 3) 데이터 생애주기로 정의되는 실무 활동의 세 가지 측면에서 논의하였다. 이와 같이 데이터 리터러시의 개념이 다양한 측면을 포괄하고 있으므로 그 개념을 여러 가지 용어로 세분화하는 것 보다는 Koltay의 의견과 같이 데이터 리터러시라는 하나의 용어로 정의하는 것이 바람직하다고 보았다. 이를 통해 다양한 데이터 관련 이해관계자들을 고려하는 유연하고 확장성 있는 접근이 가능하다고 제안하였다.

이러한 선행연구의 주장을 바탕으로 본 연구에서는 '데이터 리터러시'라는 용어를 사용하여 과학 데이터 리터러시, 연구데이터 리터러시, 데이터 정보 리터러시 등 기존의 용어와 개념들을 포괄하여 지칭하였다. 즉 데이터 리터러시는 데이터 생산자와 데이터 이용자의 관점을 반영하여 연구 활동의 원천이자 결과물인 데이터를 생성, 수집, 기술, 관리, 평가, 보존, 이용 및 재이용하는 능력을 의미한다. 이는 데이터 기반 연구 환경에서 학생 및 연구자들에게 요청되는 능력이면서 데이터를 관리하는 실무자들이 갖추어야 하는 능력으로 정의된다.

2.2 데이터 리터러시: 역량

데이터 리터러시와 관련된 세부적인 지식과 능력을 의미하는 데이터 리터러시 세부 역량(competencies)은 여러 연구들에서 정의되어 왔다. Carlson 외(2011)는 교수들과의 인터뷰 및

지리정보학(geoinformatics) 강의를 수강하는 학생들을 대상으로 다양한 데이터 리터러시 세부 역량의 중요도에 대한 인식을 조사하였다. 이와 같이 실증적인 방법론을 활용하여 도출된 12

개의 역량과 각각의 정의를 살펴보면 <표 1>과 같다. 각 역량의 용어와 제시된 순서는 상기 연구가 수행된 과제의 최종 결과로서 Nelson(2015)이 기술한 내용에 따랐다.

<표 1> 12개 데이터 리터러시 세부 역량

데이터 리터러시 세부 역량	정의
데이터 검색과 수집 (Discovery and Acquisition of Data)	- 학문분야 데이터 레포지토리를 찾아내고 활용 - 적절한 데이터 출처를 발견하고 필요할 경우 이를 가져와서 변환시켜 데이터 처리 도구에서 사용
데이터베이스 및 데이터 포맷 (Databases and Data Formats)	- 해당 학문분야에서 주로 사용되는 데이터 포맷과 유형 숙지 - 연구질문에 적합한 데이터 포맷과 유형을 파악 - 관계데이터베이스의 개념과 질의(query) 작성 방법을 이해
데이터 변환 및 상호운용성 (Data Conversion and Interoperability)	- 하나의 데이터 포맷에서 다른 포맷으로 마이그레이션 - 포맷 변경으로 인해 발생하는 위험 및 잠재적인 정보 유실 및 손상 가능성 인지 - 표준 포맷으로 데이터를 활용하는 것의 장점을 이해
데이터 처리와 분석 (Data processing and analysis)	- 해당 분야의 기본적인 데이터 처리 및 분석 도구를 숙지 - 이러한 도구들이 데이터에 미치는 영향 이해 - 반복적인 데이터 분석을 자동화하는 적절한 업무흐름 관리 도구를 활용
데이터 시각화 및 표현 (Data visualization and representation)	- 해당 분야의 기본적인 시각화 도구를 능숙하게 사용 - 데이터를 표, 도표, 다이어그램으로 제시할 때 오해의 소지가 있고 불분명한 표현을 피함 - 데이터를 보여줄 때 시각화의 목적에 대한 이해를 바탕으로 지도, 그래프, 애니메이션, 비디오 등 다양한 시각화 유형 중 적절한 것을 선택
데이터 관리와 조직 (Data Management and Organization)	- 데이터 생애주기를 이해하고 데이터관리계획을 개발 - 원본 데이터와 하위 데이터 혹은 처리된 데이터 간 관계성을 유지 - 데이터 관리와 도큐멘테이션에 대한 표준 운영 절차 작성
데이터 품질 및 도큐멘테이션 (Data quality and documentation)	- 데이터의 명백한 인위성(artifacts), 불완전성 또는 손상을 인지, 기록하고 해결 - 데이터세트의 잠재적인 문제 파악이 가능하도록 메타데이터 활용 - 다른 사람들이 연구결과를 재현할 수 있도록 데이터에 대한 충분한 도큐멘테이션 제공 - 데이터 출처를 표시하고 데이터세트의 버전 정보를 명시
메타데이터와 데이터 기술 (Metadata and data description)	- 메타데이터의 필요성을 이해하고 누구나 데이터를 이해하고 활용할 수 있게 데이터를 기술 - 외부 정보원으로부터 받은 메타데이터를 읽고 해석하는 능력 개발 - 데이터 공유를 촉진하는 온톨로지의 구조와 목적 이해
데이터 관리 및 공유 문화 (Cultures of practices)	- 데이터의 관리, 공유, 큐레이션 및 보존과 관련된 해당 분야의 실무, 가치 및 규범을 인식 - 메타데이터, 데이터 품질, 포맷 등 해당 분야에서 적절한 데이터 표준을 인지하고 어떻게 이를 적용할 것인지를 이해
윤리 및 데이터 인용 (Ethics and attribution)	- 지적재산권, 개인정보보호, 비밀유지 관련 이슈 이해 - 데이터 공유와 관련된 해당 분야에서의 관습 이해 - 외부 정보원으로부터 온 데이터에 대한 적절한 인용 제시
데이터 큐레이션 및 재이용 (Data curation and reuse)	- 데이터는 원래 생산된 목적 이상의 가치를 가질 수 있으며 연구결과 검증이나 다른 목적으로 이용될 수 있음을 이해 - 데이터세트의 어떤 요소가 미래 가치를 가질 수 있을지 식별 - 데이터 큐레이션은 복잡하고 많은 비용이 요구되지만 필수불가결한 활동임을 이해 - 데이터는 생성 단계에서부터 생애주기를 거쳐 궁극적인 큐레이션에 이르기까지 관리되어야 함을 인지 - 데이터 큐레이션에 필요한 계획과 활동을 명시 - 데이터를 인용하는 방법과 데이터가 인용될 수 있게 하는 방법 이해
데이터 보존 (Data preservation)	- 데이터 보존의 장점과 비용을 인식 - 데이터 보존에 필요한 기술, 자원 및 조직 차원의 구성 요소 이해 - 데이터의 가치와 재생산성에 부합하는 보존 모범실무 활용

출처: Nelson, M. S. (2015). Where do we go from here? Furthering developing the data information literacy competencies. In Carlson, J., & Johnston, L. R. (Eds.), *Data information literacy: Librarians, data, and the education of a new generation of researchers* (pp.231-245). Purdue University Press.

또 다른 연구로 Prado와 Marzal(2013)은 기존의 정보 리터러시 표준을 바탕으로 데이터 리터러시 세부 역량을 도출하였다. 이를 위해 미국대학도서관협회(Association of College and Research Libraries, ACRL)에서 제정한 고등교육에서의 정보 리터러시 표준과 함께 분야별 요구를 반영하여 제정된 과학·공학, 인류학·사회학, 정치학, 언론학 분야에서의 정보 리터러시 표준 및 가이드라인, 그리고 미국학교도서관사서협회(American Association of School

Librarians, AASL)의 정보 리터러시 표준을 참고하였다. 이렇게 도출된 데이터 리터러시 세부 역량을 5개 영역인 1) 데이터 이해; 2) 데이터 검색 및 수집; 3) 데이터 해석 및 평가; 4) 데이터 관리; 5) 데이터 이용으로 구분하여 각 영역 아래 세부 역량과 정의를 제시하였다(〈표 2〉 참조). 저자들은 이러한 핵심 역량을 바탕으로 데이터 리터러시 교육 프로그램 개발 시 필요한 내용을 포괄할 수 있다고 보았다. 기존의 문헌들을 종합하여 데이터 리터러시

〈표 2〉 데이터 리터러시 핵심 역량

영역	세부영역	역량	내용
데이터 이해 (Understanding data)	데이터 개념 (What is data)	데이터의 개념과 다양한 데이터 유형을 이해	데이터의 정의, 데이터의 유형(데이터의 기원, 포맷, 이용 라이선스에 따라 달라짐)
	사회에서의 데이터 (Data in society)	사회에서 데이터의 역할, 데이터 생성방법과 생성주체, 적용 가능 분야 및 데이터 이용의 영향을 주지	데이터 생산자와 이용자, 데이터 생애주기, 데이터 적용, 과학과 사회에 대한 데이터의 영향, 데이터 재이용에 영향을 미치는 저작권과 라이선스
데이터 검색과 확보 (Finding and/or obtaining data)	데이터 출처 (Data source)	가능성 있는 데이터 출처들을 인식하고 이들을 평가하여 정보 요구나 주어진 문제에 가장 적합한 것을 선택	데이터 출처, 데이터 출처를 평가하는 기준
	데이터 확보 (Obtaining data)	기존의 데이터로 주어진 문제나 요구를 해결할 수 없는 경우를 인지하고 새로운 데이터를 확보하기 위한 조사나 연구를 수행	원 데이터를 수집하는데 필요한 주요 연구 방법론
데이터 해석과 평가 (Reading, interpreting and evaluating data)	데이터 해석 (Reading and interpreting data)	문자, 숫자, 그래프 등 데이터의 다양한 형태와 각각의 특징들을 주지하고 이를 해석	데이터를 제시하고 표현하는 방법
	데이터 평가 (Evaluating data)	데이터를 비판적으로 평가	데이터 평가 기준(저자, 데이터 수집 및 분석 방법, 비교가능성, 추론, 데이터 요약 포함)
데이터 관리 (Managing data)	데이터 및 메타데이터 수집과 관리 (Data and metadata collection and management)	데이터 식별, 관리 및 재이용을 위해 데이터와 관련된 메타데이터를 유지하는 필요성을 이해	메타데이터, 서지관리도구, 데이터베이스, 데이터 관리 레포지터리, 정책 및 실무
데이터 이용 (Using data)	데이터 취급 (Data handling)	데이터 분석을 위한 준비 및 결과를 얻기 위한 데이터 분석, 필요한 분석 도구의 이용 방법 인지	데이터 변환, 데이터 분석 도구 취급
	데이터 분석 및 표현 (Producing elements for data synthesis)	데이터의 특성, 목적 및 연구문제에서 목표포 함은 이용자층에 부합하는 방식으로 데이터를 종합하고 그 결과를 제시	적절한 데이터 표현 방법(표, 그래프 등) 선택, 분석도구나 독립형 어플리케이션의 취급
	데이터의 윤리적 이용 (Ethical use of data)	데이터의 윤리적 이용, 다른 사람으로부터 확보한 데이터의 출처 명시, 사용된 방법론과 결과 해석을 투명하고 정직하게 기술	데이터의 윤리적 이용에 대한 개념, 데이터 인용 방법

출처: Prado, J. C., & Marzal, M. Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*, 63(2), 130-131.

세부 역량을 제시한 또 다른 연구로 Ridsdale, Rothwell, Smit, Ali-Hassan, Bliemel, Irvine, Kelley, Matwin, Wuetherick, & Bradley(2015)는 데이터 리터러시의 5가지 영역을 1) 개념적 틀(conceptual framework); 2) 데이터 수집; 3) 데이터 관리; 4) 데이터 평가; 5) 데이터 적용(data application)으로 제시하였다. 또한 데이터 리터러시 세부 역량을 세 가지 수준인 개념적(conceptual) 역량, 핵심(core) 역량, 고급

(advanced) 역량으로 구분하였다. 저자들은 문헌연구를 바탕으로 도출된 22개의 역량을 각 영역과 역량 수준으로 구분한 데이터 리터러시 세부 역량 매트릭스(matrix)를 제안하였다(〈표 3〉 참조). 이와 같은 역량 매트릭스는 데이터 리터러시 관련 지식과 능력을 세분화하였다는 점에서 의의가 있다. 그러나 각 역량 및 역량수준에 대한 정의를 제시하지 않고 있어 역량수준 간 명확한 구분이 어렵다는 한계가 있다.

〈표 3〉 데이터 리터러시 세부 역량 매트릭스

역량수준 \ 영역	개념적 틀	데이터 수집	데이터 관리	데이터 평가	데이터 적용
개념적 역량 (Conceptual competencies)	데이터 입문 (Introduction to Data)			- 데이터 도구 (Data tools)	- 비판적 사고 (Critical thinking) - 데이터 문화 (Data culture) - 데이터 윤리 (Data ethics)
핵심 역량 (Core competencies)		- 데이터 검색 및 수집 (Data discovery and collection) - 데이터 및 출처의 품질 평가와 확인 (Evaluating and ensuring quality of data and sources)	- 데이터 조직 (Data organization) - 데이터 조작 (Data manipulation)	- 기초 데이터 분석 (Basic data analysis) - 데이터 해석 (Data interpretation (understanding data)) - 데이터 이용 시 문제점 파악 (Identifying problems using data) - 데이터 시각화 (Data visualization) - 데이터 분석 결과 제시 (Presenting data (verbally)) - 데이터 기반 의사 결정 (Data Driven Decision Making (DDDM))	- 데이터 인용 (Data citation) - 데이터 공유 (Data sharing)
고급 역량 (Advanced competencies)			- 데이터 변환 (Data conversion (from format to format)) - 메타데이터 생성 및 이용 (Metadata creation and use) - 데이터 큐레이션 및 재이용 (Data curation and reuse) - 데이터 보존 (Data preservation)		- 데이터 기반 결정 및 결론 평가 (Evaluating Decisions/Conclusions Based on Data)

출처: Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., ... & Wuetherick, B. (2015). Strategies and best practices for data literacy education: Knowledge synthesis report, Dalhousie University. p.38.

상기 기술한 연구들에서 제시된 데이터 리터러시 세부 역량을 비교한 결과는 <표 4>와 같다. Carlson 외(2011)와 Nelson(2015)이 제시한 역량들을 중심으로 살펴보았을 때 모든 연구들에서 제시된 역량은 1) 데이터 검색과 수집; 2) 데이터 처리와 분석; 3) 데이터 관리와 구성; 4) 메타데이터와 데이터 기술; 5) 윤리 및 데이터 인용인 것으로 나타났다. Prado와

Marzal(2013)의 연구에서는 제시되지 않았으나 나머지 두 연구에서는 제안된 역량으로는 1) 데이터 변환 및 상호운용성; 2) 데이터 시각화 및 표현; 3) 데이터 품질 및 도큐멘테이션; 4) 데이터 관리 및 공유 문화; 5) 데이터 큐레이션 및 재이용; 6) 데이터 보존이 있다. 이 외에도 데이터베이스 및 데이터 포맷 관련 역량은 Carlson 외(2011)와 Nelson(2015)에서만 언급되었다.

<표 4> 데이터 리터러시 세부 역량 비교

Carlson 외(2011): Nelson(2015)	Prado and Marzal(2013)		Ridsdale, Rothwell, Smit 외(2015)	
데이터 검색과 수집	데이터 검색과 수집	- 데이터 출처 - 데이터 수집	데이터 수집	- 데이터 검색 및 수집
데이터베이스 및 데이터 포맷				
데이터 변환 및 상호운용성			데이터 관리	- 데이터 조작 - 데이터 변환
데이터 처리와 분석	데이터 이용	- 데이터 취급 - 데이터 분석 결과 제시	데이터 평가	- 기초 데이터 분석 - 데이터 해석 - 데이터 이용 시 문제점 파악 - 데이터 분석 결과 제시 - 데이터 기반 의사 결정 - 데이터 기반 결정 및 결론 평가
	데이터 해석 및 평가	- 데이터 해석 - 데이터 평가	데이터 평가	- 데이터 도구
데이터 시각화 및 표현			데이터 적용	- 비판적 사고
			데이터 평가	- 데이터 시각화
데이터 관리와 조직	데이터 관리	- 데이터, 메타데이터 수집 및 관리	데이터 관리	- 데이터 조직
	데이터 이해	- 데이터 개념 - 사회에서의 데이터	개념적 틀	- 데이터 입문
데이터 품질 및 도큐멘테이션			데이터 수집	- 데이터 및 출처의 품질 평가와 확인
메타데이터와 데이터 기술	데이터 관리	- 데이터, 메타데이터 수집 및 관리	데이터 관리	- 메타데이터 생성 및 이용
데이터 관리 및 공유 문화			데이터 적용	- 데이터 문화
윤리 및 데이터 인용	데이터 이용	- 데이터의 윤리적 이용	데이터 적용	- 데이터 윤리 - 데이터 인용 - 데이터 공유
데이터 큐레이션 및 재이용			데이터 관리	- 데이터 큐레이션 및 재이용
데이터 보존			데이터 관리	- 데이터 보존

이를 바탕으로 본 연구에서는 Carlson 외(2011)와 Nelson(2015)이 제시한 12개 데이터 리터러시 세부 역량을 교육 프로그램을 분석을 위한 코드로 활용하였다. 구체적인 분석대상과 방법은 다음 절에서 기술하였다.

3. 대학도서관 연구데이터 관리 교육 프로그램 분석

3.1 분석대상 및 방법

본 연구에서는 북미에서 연구데이터 관리 교육 프로그램을 제공하는 51개 대학을 조사대상으로 하였다. 이들 기관을 선정하기 위해 북미의 연구중심 대학 또는 ARL 회원도서관이 속하는 대학을 포함한 138개 기관 중에서 연구데이터 관리 서비스를 제공하고 있는 121개 대학을 확인하였다. 이 중에서 연구데이터 관리 웹사이트 또는 도서관 가이드를 통해 교육 프로그램에 대한 메뉴를 제공하고 있는 51개 기관을 조사대상으로 삼았다.

먼저 이들 51개 기관의 교육 프로그램 메뉴에서 명시하고 있는 워크숍이나 세미나 등 교육 프로그램의 제공 방식과 수 및 제공 일자를 살펴보았다. 첫째, 조사대상 기관의 교육 프로그램 제공 방식은 크게 집합 교육과 온라인 교육으로 구분할 수 있었다. 집합 교육을 제공하는 기관은 51개 기관 중 43개 기관(84.3%)으로 다수를 차지하고 있었다. 나머지 8개 대학 중 5곳은 집합 교육 없이 자체 개발한 온라인 튜토리얼을 제공하고 있었으며 3곳은 집합 교육 없이 외부 기관에서 개발된 온라인 튜토리

얼의 링크를 제공하고 있었다. 집합 교육을 제공하는 43개 기관 중에서도 9개 기관은 외부 기관에서 개발된 온라인 튜토리얼을 함께 제공하고 있는 것으로 나타났다.

둘째, 교육 프로그램의 수는 집합 교육을 제공하고 있는 43개 기관에만 한정하여 살펴보았다. 온라인 튜토리얼의 경우 주로 하나의 튜토리얼이 여러 개의 모듈로 구성되어 연구데이터 관리에 대한 다양한 주제를 다루고 있어 교육 프로그램의 수를 특정하기 어려운 한계가 있었기 때문에 제외하였다. 집합 교육을 제공하는 기관에서 명시하고 있는 워크숍 및 세미나 등 교육 프로그램의 수는 총 317개이었으며 기관별로 최소 1개에서부터 최대 32개에 이르기까지 기관 간 편차가 큰 것으로 나타났다([부록 1] 참조). 또한 이들 43개 기관에서 제공하는 교육 프로그램 수의 평균은 7.4개, 중간값은 4개인 것으로 나타났다.

셋째, 교육 프로그램의 제공 일자와 관련하여 51개 기관 중 33개 기관에서는 연구데이터 관리 교육 프로그램 메뉴에서 프로그램의 제공 일자를 명시하고 있었다. 이들 33개 기관 중 21개 기관에서는 현재 학기, 즉 2018년 9월에서 11월 사이에 제공되는 교육 프로그램을 안내하고 있었으며, 11개 기관은 과거에 개설되었던 교육 프로그램을, 1개 기관에서는 과거에서부터 현재까지의 교육 프로그램에 대한 설명을 제시하고 있었다. 제공 일자를 명시하지 않은 18개 기관 중에서 7개 기관은 집합 교육만을 실시하고 있지만 제공 일자는 언급하지 않았으며 3개 기관은 집합 교육과 외부 온라인 튜토리얼에 대한 안내를 둘 다 제공하는 기관이지만 역시 제공 일자에 대한 설명은 없었다. 그 외 8개 기관은 집합 교육 없이 온라인 튜토리얼만을

제공하는 기관이어서 언제든 접근할 수 있는 온라인 교육의 특성상 제공 일자를 명시하지 않고 있었다([부록 1] 참조).

이와 더불어 본 연구에서는 교육 프로그램의 내용 분석을 위해 집합 교육을 제공하는 43개 기관에서 명시하고 있는 317개 교육 프로그램의 제목을 수집하였다. 또한 집합 교육 없이 온라인 튜토리얼만을 제공하는 8개 기관 중에서 자체적으로 개발한 온라인 튜토리얼을 제공하는 5개 기관을 중심으로 각각의 튜토리얼 목차에서 가장 상위 수준의 제목 42개를 수집하였다. 이와 같이 수집된 제목들을 바탕으로 선행 연구에서 공통적으로 제시된 11개 데이터 리터러시 세부 역량과 공통적으로 제시되지는 않았으나 코딩 과정에서 관련 교육 프로그램이 존재하였으므로 포함된 데이터베이스 및 데이터 포맷 역량 등 12개 역량에 따라 코딩을 실시하였다. 코딩 과정에서는 각각의 교육 프로그램에 대한 개요를 참고하여 가장 중점적으로 다

루고 있는 데이터 리터러시 세부 역량에 초점을 맞추어 분석을 수행하였다. 이를 통해 북미 대학도서관의 연구데이터 관리 교육 프로그램에서 중점적으로 다루어지고 있는 데이터 리터러시 세부 역량과 그렇지 않은 역량을 구분하여 살펴볼 수 있었다. 분석된 교육 프로그램의 내용은 다음 절에서 프로그램 제공 방식에 따라 구분하여 역량별로 서술하였다.

3.2 교육 프로그램 내용

3.2.1 집합 교육

집합 교육 방식으로 제공되는 교육 프로그램의 경우 12개 데이터 리터러시 세부 역량 중 일부에 집중되는 경향을 보이고 있었다. <표 5>에서는 각각의 데이터 리터러시 세부 역량별로 집합 교육 프로그램을 제공하고 있는 기관 수와 프로그램 수를 제시하고 있다. 각각의 역량에 따른 교육 프로그램 내용은 다음과 같다.

<표 5> 데이터 리터러시 세부 역량별 집합 교육 프로그램 수 및 프로그램 제공 기관 수

데이터 리터러시 역량	기관 수	기관 %†	프로그램 수	프로그램 %
데이터 검색 및 수집	9	20.9	16	5.0
데이터베이스 및 데이터 포맷	3	7.0	3	0.9
데이터 변환 및 상호운용성	7	16.3	9	2.8
데이터 처리 및 분석	24	55.8	140	44.2
데이터 시각화 및 표현	15	34.9	34	10.7
데이터 관리 및 조직	33	76.7	80	25.2
데이터 품질 및 도큐멘테이션	4	9.3	4	1.3
메타데이터 및 데이터 기술	4	9.3	4	1.3
데이터 관리 공유 문화	6	14.0	10	3.2
윤리 및 데이터 인용	8	18.6	11	3.5
데이터 큐레이션 및 재이용	2	4.7	4	1.3
데이터 보존	2	4.7	2	0.6
합계			317	100.0

† : 집합 교육을 제공하는 43개 기관 중 해당 리터러시 역량과 관련된 교육 프로그램을 제공하는 기관 비율

1) 데이터 검색 및 수집

집합 교육을 제공하는 조사 대상 기관 43곳 중에서 데이터 검색 및 수집 역량에 관련된 교육 프로그램을 제공하는 기관은 모두 9곳이었다. 이들 기관이 제공하는 교육 프로그램의 수는 1개 기관에서 5개, 다른 1개 기관에서 4개, 그리고 나머지 7개 기관에서는 각각 1개의 교육 프로그램을 제공하고 있는 것으로 나타났다. 따라서 이들 8개 기관에서 제공하는 교육 프로그램의 총 수는 16개이었으며 내용별로 살펴보면 <표 6>과 같다.

웹 사이트에서 원하는 위치의 데이터를 자동으로 추출하는 기법인 웹 스크래핑에 대한 교육 프로그램은 3개인 것으로 나타났고, 3개 기관에서 각각 하나씩을 제공하고 있었다. 또한 다양한 데이터의 유형 및 성격에 따라 검색 및 수집 방법을 논의하는 교육 프로그램들이 운영되고 있었다. 이와 관련하여 트위터 데이터와 같은 특정 소셜미디어 플랫폼에서 추출된 데이터 수집에서부터 광범위한 의미를 가지는 오픈 데이터, 통계 정보원 등을 검색 및 수집 대상으로 하는 프로그램도 존재하였다. 데이터 수집을 위한

설문지 설계 방법과 관련된 워크숍도 있었으며 1개 기관에서 관련 프로그램을 2개 개설하고 있었다. 이 외에도 사회과학 분야의 대표적인 데이터 레포지터리인 ICPSR(Inter-university Consortium for Political and Social Research)에 대한 프로그램은 2개 기관에서 각각 하나씩 제공되고 있었으며 적합한 데이터 레포지터리를 선택하는 방법과 관련된 워크숍도 운영되고 있었다.

2) 데이터베이스 및 데이터 포맷

데이터베이스 및 데이터 포맷과 관련된 교육 프로그램을 제공하는 기관은 43개 기관 중 3개 대학에 불과하였고 기관별로 프로그램 1개씩을 제공하고 있었다. 관계 데이터베이스와 SQL 질의문에 대한 워크숍을 제공하고 있는 대학은 캘리포니아주립대(데이비스)와 스탠포드대 2개 기관이었다. 데이터 포맷에 대한 교육으로 노스캐롤라이나주립대에서는 Application Programming Interfaces(APIs)에서 가장 많이 활용되는 데이터 포맷인 JSON에 대한 워크숍을 운영하고 있었다.

<표 6> 데이터 검색 및 수집 역량 관련 교육 프로그램

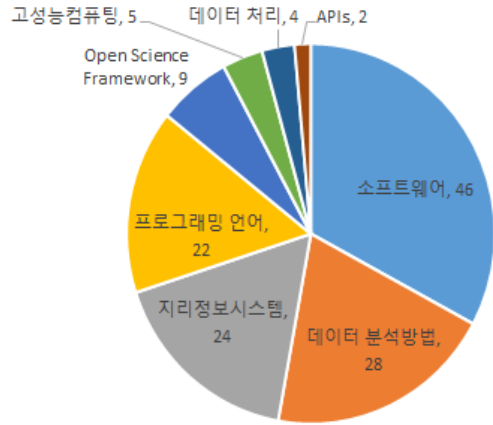
교육 프로그램 내용		프로그램 수	프로그램 제공 기관	
데이터 검색 및 수집	웹 스크래핑(Web scraping)	3	조지메이슨대, 콜로라도대, 예일대	
	설문지 설계 방법	2	조지아 주립대	
	데이터 유형별	트위터 데이터	1	조지메이슨대
		마케팅 데이터	1	조지아주립대
		오픈 데이터	1	캘리포니아대-리버사이드
		위치 데이터	1	캘리포니아대-데이비스
		인구 데이터	1	조지메이슨대
		통계 정보원	1	조지메이슨대
레포지터리	ICPSR	2	조지메이슨대, 조지아주립대	
	데이터 레포지터리 선택	3	플로리다대, 존스홉킨스대, 듀크대	
합계		16		

3) 데이터 변환 및 상호운용성

데이터 변환 및 상호운용성과 관련된 교육 프로그램은 주로 데이터 정제(data cleaning)와 관련된 것으로 7개 기관에서 9개 프로그램을 제공하고 있었다(〈표 7〉 참조). 7개 기관 중 2곳에서 각각 2개의 프로그램을, 나머지 5개 기관은 하나의 프로그램을 운영하고 있었다. 제공되는 9개 프로그램 중 7개는 OpenRefine이라는 어플리케이션을 활용하여 데이터를 정제하는 방법을 설명하고 있었다. 그 외에도 프로그래밍 언어인 R을 활용한 데이터 정제 기법에 대한 교육 프로그램과 텍스트데이터의 정제에 초점을 맞춘 프로그램이 각각 1개씩 제공되고 있었다.

4) 데이터 처리 및 분석

데이터 처리 및 분석 역량은 43개 기관 중 절반이 넘는 24개 기관에서 140개의 관련 교육 프로그램을 제공하는 것으로 나타났으며 이는 북미 연구데이터 관리 교육 프로그램에서 가장 많이 다루어지고 있는 영역임을 알 수 있다. 이들 교육 프로그램은 주로 데이터 분석을 위한 소프트웨어 활용, 데이터 분석 방법, 프로그래밍 언어, 지리정보시스템에 대한 것이었으며 이를 정리하면 〈그림 2〉과 같다.



〈그림 2〉 데이터 처리 및 분석 역량 관련 교육 프로그램 세부 영역

데이터 분석을 위한 소프트웨어 교육 프로그램은 46개로 가장 높은 비중을 차지하고 있으며 14개 대학에서 제공되고 있었다. 〈표 8〉에서와 같이 통계 소프트웨어에 대한 워크숍이 14개로 가장 많았고 6개 대학에서 운영되고 있었다. 또한 질적 데이터 분석 소프트웨어인 NVivo에 대한 교육 프로그램은 2개 기관에서 각각 2개씩 모두 4개인 것으로 나타났다. 이들 소프트웨어는 정량적 또는 정성적 데이터 분석에 있어 가장 널리 사용되는 분석도구인 만큼 이에 대한 교육 프로그램의 운영 빈도가 높다는 것을 알 수 있다.

〈표 7〉 데이터 변환 및 상호운용성 역량 관련 교육 프로그램

교육 프로그램 내용		프로그램 수	프로그램 제공 기관
데이터 정제	OpenRefine	7	보스턴대, 브라운대, 코넬대, 조지메이슨대, 노스캐롤라이나주립대, 스탠포드대, 위스콘신대-매디슨
	R	1	노스캐롤라이나주립대
	텍스트데이터 정제	1	브라운대
합계		9	

〈표 8〉 데이터 처리 및 분석: 소프트웨어 교육 프로그램

교육 프로그램 내용		프로그램 수	프로그램 제공 기관
소프트웨어	통계 소프트웨어(SPSS, SAS, Stata)	14	보스턴대, 듀크대, 조지메이슨대, 조지아주립대, 럿거스대, 텍사스대-알링턴
	Nvivo	4	조지메이슨대, 조지아주립대
	Jupyter Notebook	4	보스턴대, 하버드대, 노스캐롤라이나주립대, 라이스대
	Rstudio	4	듀크대, 노스캐롤라이나주립대
	Excel	3	조지메이슨대, 라이스대
	MATLAB	3	보스턴대
	Pandas	2	노트르담대, 텍사스대-알링턴
	GitLab	1	듀크대
	ImageJ	1	보스턴대
	Shell	1	보스턴대
	Pandoc	1	보스턴대
	Tropy	1	콜로라도대
	Voyant tools, AntConc	1	노스캐롤라이나주립대
	기타	The Carpentries	5
Library Carpentry		1	캘리포니아대-샌디에고
합계		46	

이와 더불어 다양한 프로그래밍 언어를 지원하여 데이터 분석을 가능하게 하는 오픈 소스 소프트웨어인 Jupyter Notebook에 대한 교육 프로그램은 4개 기관에서 각각 하나씩 운영되고 있었다. 통계 분석 및 데이터 시각화를 위한 프로그래밍 언어인 R의 활용을 지원하는 RStudio 관련 교육은 2개 기관에서 4개 프로그램을 제공하고 있었다. 엑셀을 활용한 데이터 분석 방법에 대한 교육의 경우 2개 기관에서 3개의 워크숍이 운영되고 있었으며, 수치 해석을 위한 공학용 소프트웨어인 MATLAB 관련 교육 프로그램은 보스턴대 1개 기관에서 3개의 워크숍으로 제공되고 있었다. 이 외에도 Pandas, ImageJ, Pandocs 등 다양한 데이터 분석 소프트웨어에 대한 교육 프로그램들이 존재하였다.

기타에 포함된 6개 프로그램 중 4개는 Software

Carpentry라는 외부 교육기관에 의해 제공되는 워크숍이었으며 1개는 Data Carpentry, 다른 하나는 Library Carpentry에 의해 제공되는 워크숍이었다. 2018년 현재 Software Carpentry와 Data Carpentry가 병합되어 The Carpentries라는 명칭으로 운영되고 있으며 10개국 70개 기관이 회원으로 가입되어 있다(The Carpentries, n.d.). Library Carpentry는 The Carpentries의 교육 모형과 운영 방식을 준용하면서 2015년 조직된 단체로 사서 및 정보전문가들을 대상으로 업무에서 필요한 소프트웨어 활용 능력과 데이터 관리 및 분석 지식을 교육하는 것을 목적으로 한다. 이를 통해 궁극적으로는 사서들이 소속 기관의 연구자 또는 학생들을 대상으로 연구데이터 관리 및 소프트웨어 활용 교육을 제공할 수 있도록 하는 것을 지향한다(Library

Carpentry, n.d.).

Teal, Cranston, Lapp, White, Wilson, Ram, & Pawlik(2015)에 따르면 1998년 시작된 Software Carpentry는 보통 이를 동안 운영되는 실습 중심의 집중 워크숍(2-days intensive hands-on workshops)을 통해 연구자들이 필요로 하는 소프트웨어 활용법과 프로그래밍 언어를 교육하는 비영리 기관이다. 워크숍은 자원 봉사 강사들에 의해 운영되며 교육 내용은 소프트웨어 교육이나 오픈 소스 소프트웨어 등 관련 커뮤니티에 속한 사람들의 협력을 통해 개발된다. Teal 외는 이러한 교육 방식을 커뮤니티와 함께 하는 실습 중심 워크숍 모형(a hands-on workshop model with community)이라고 명명하였다. 고품질의 교육 자료와 강사와의 상호작용을 바탕으로 하는 실습 중심 워크숍이 소프트웨어 활용 능력을 향상시키고자 하는 연구자들의 요구에 효과적으로 부합한다고 보았다.

Data Carpentry는 2014년에 Software Carpentry의 자매기관으로 시작되었으며 연구데이터를 생산하는 모든 연구자들을 대상으로 데이터 관리 및 분석 능력을 교육하는 것에 중점을 두고 있다. 이들의 비전은 보편적인 데이터 리터러시를 교육하는 커뮤니티를 구성하는 것이며 세 가지의 사명을 명시하고 있는데 1) 효율적이고 공유할 수 있으며 재현 가능한 연구 수행을 위한 핵심 데이터 능력 교육; 2) 모두에게 공개되면서 품질이 우수하고 특정 분야의 요구에 부합하는 교육 제공; 3) 활발히 활동하면서 포괄적이고 다양한 강사 커뮤니티 양성을 통해 재현 가능한 연구를 커뮤니티 규범으로 장려하고 실천하는 것이다(Data Carpentry, n.d.).

The Carpentries에서는 연구자들에게 필요

한 컴퓨팅 능력에 대한 워크숍을 기관의 요청에 의해 제공하며 기관에서는 이에 대한 비용을 지불한다. 일반적으로 프로그래밍 언어인 R이나 Python 관련 교육, Unix Shell, MATLAB 등 소프트웨어 교육 및 데이터 처리, 분석 또는 시각화에 관한 교육 등을 수행하고 있다(The Carpentries, n.d.). 본 연구의 조사대상 기관 중에서는 스탠포드대, 텍사스대(알링턴), 캘리포니아대(리버사이드), 앨버타대 등 4개 대학에서 이 기관을 통해 관련 교육을 제공하고 있었다. 스탠포드대의 경우 Python, Git, Shell에 관한 워크숍으로 명시하고 있으나 나머지 3개 대학에서는 교육 내용에 대한 언급은 없었다.

Library Carpentry 워크숍을 명시한 기관은 캘리포니아대(샌디에고) 1곳이었으며 사서들을 대상으로 하는 교육 프로그램이라는 점에서 본 연구의 분석대상인 연구자 대상 교육 프로그램과는 차이가 있다. 그러나 해당 대학의 교육 프로그램 메뉴에 이 워크숍이 포함되어 있었고 데이터 관리 및 분석을 위한 컴퓨팅 기술 현황과 소프트웨어 활용법을 교육한다는 점과 사서들의 데이터 리터러시 함양을 위한 계속교육 프로그램이라는 점에서 의미를 가진다고 판단하여 분석대상에 포함하였다.

다음으로 데이터 분석 방법에 대한 교육 프로그램은 12개 기관에서 28개가 제공되고 있었다(〈표 9〉 참조). 그 중에서도 10개 프로그램이 텍스트데이터 분석 방법에 대한 것이었는데 보스턴대 2개, 브라운대 1개, 노스캐롤라이주립대 3개, 콜로라도대에서 4개의 관련 프로그램이 제공되고 있었다. 이들 프로그램에서는 XML 인코딩 기반 분석, 토픽 모델링 기법, R이나 Sketch Engine 등 특정 프로그래밍 언어 또는 소프트

〈표 9〉 데이터 처리 및 분석: 데이터 분석 방법 교육 프로그램

교육 프로그램 내용		프로그램 수	프로그램 제공 기관
데이터 분석 방법	텍스트데이터 분석	10	보스턴대, 브라운대, 노스캐롤라이나주립대, 콜로라도대
	통계적 분석	5	컬럼비아대, 콜로라도대
	네트워크 분석	4	존스홉킨스대, 콜로라도대, 예일대
	정성적 분석	2	조지메이슨대, 조지아주립대
	공간데이터 분석	1	듀크대
	차세대 염기서열 데이터 분석	1	하버드대
	오디오 데이터 분석	1	노스캐롤라이나주립대
	원격 탐사 데이터 분석	1	캘리포니아대-데이비스
기타	3	토론토대	
합계		28	

웨어를 활용한 분석 방법을 다루고 있었다. 통계적인 분석 기법에 대한 교육 프로그램은 2개 대학에서 5개 프로그램이, 네트워크 분석과 관련해서는 3개 대학에서 4개의 워크숍이 제공되고 있었다. 정성적 분석 방법에 대한 교육 프로그램은 2개 기관에서 각각 1개씩 운영되고 있었다. 그 외 공간데이터 분석이나 염기서열 데이터 분석 등 특정 유형의 데이터 분석 방법에 대한 교육 프로그램도 존재하였다. 기타 범주에 포함된 3개 프로그램들은 데이터 분석 방법에 대한 개괄적인 내용을 다루는 워크숍이었다.

지리정보시스템(Geographical Information

System, GIS)과 관련 소프트웨어에 대한 교육 프로그램은 12개 기관에서 24개가 제공되고 있었다(〈표 10〉 참조). 이 중에서 ArcGIS라는 지리 데이터 분석 소프트웨어에 대한 교육 프로그램이 12개로 가장 많았으며 7개 기관에서 운영하고 있었다. QGIS라는 오픈 소스 기반 소프트웨어에 대한 교육 프로그램은 5개 기관에서 각각 하나씩 제공되고 있었으며, GIS에 대한 개괄적인 설명을 제시하는 입문 프로그램 역시 4개 기관에서 하나씩 운영되고 있었다. 기타 범주에 포함되는 교육 프로그램들은 GIS를 기반으로 고지도를 웹 지도로 구축하는 방법이나 지리 데이터의 코딩 방법 등을 다루고 있었다.

〈표 10〉 데이터 처리 및 분석: 지리정보시스템 교육 프로그램

교육 프로그램 내용		프로그램 수	프로그램 제공 기관
지리정보시스템	ArcGIS	12	코넬대, 듀크대, 조지메이슨대, 조지아주립대, 존스홉킨스대, 텍사스대-알링턴, 예일대
	QGIS	5	컬럼비아대, 코넬대, 듀크대, 조지메이슨대, 캘리포니아대-데이비스
	GIS 입문	4	보스턴대, 캘리포니아대-데이비스, 캘리포니아대-리버사이드, 콜로라도대
	기타	3	조지메이슨대, 조지아주립대, 노스캐롤라이나주립대
합계		24	

프로그래밍 언어에 대한 교육 프로그램은 10개 대학에서 22개가 운영되고 있었으며 그 중 9개 프로그램은 보스턴대 1곳에서 제공되고 있었다. <표 11>에서 나타나듯이 R을 교육하는 프로그램 수가 22개 중 10개로 가장 많은 비중을 차지했으며 Python 교육 프로그램이 6개로 그 다음을 차지하였다. 8개 기관에서 R 관련 워크숍을, 4개 기관에서 Python 워크숍을 제공하고 있었다. 나머지 프로그래밍 언어는 Scalar를 제외하고 보스턴대에서 교육하고 있는 것으로 나타났다. 따라서 대부분의 프로그래밍 언어 교육 프로그램은 R이나 Python을 중심으로 운영되고 있음을 알 수 있다.

데이터 처리 및 분석과 관련되는 기타 교육 내용으로 Open Science Framework(OSF) 이 있으며 7개 기관에서 9개의 관련 워크숍이 제공되고 있는 것으로 나타났다. OSF는 연구 생애 주기 동안 발생하는 파일, 데이터, 코드, 프로토콜 등 모든 산출물을 관리할 수 있으며 접근 통제 및 업무흐름 관리를 가능하게 하는 오픈 소스 기반 레포지터리이다(Center for Open Science, n.d). 이는 연구 공개 및 재현성(reproducibility)

을 지원하며 연구자들의 데이터 관리를 보조하는 플랫폼으로 소개되고 있었다. 또한 고성능 컴퓨팅(high performance computing)과 관련된 워크숍도 5개 기관에서 하나씩 운영되고 있었다. 데이터 처리와 관련된 교육은 3개 기관에서 4개의 워크숍으로 제공되고 있었으며 데이터를 분석하기 좋은 형태로 정리(tidy data)하는 방법에 대해 안내하고 있었다. 그 외에 OpenMP와 Google Earth Engine과 같은 APIs에 대한 교육도 2개 기관에서 하나씩 운영되는 것으로 나타났다.

5) 데이터 시각화 및 표현

데이터 시각화 및 표현과 관련된 교육 프로그램은 15개 기관에서 34개가 운영되고 있었으며 이 중 23개 프로그램은 특정 소프트웨어나 프로그래밍 언어를 활용한 시각화 방법에 대해 교육하고 있었다(<표 12> 참조). 그 중에서도 가장 많은 교육 프로그램에서 다루고 있는 소프트웨어는 인터랙티브 데이터 시각화를 가능하게 하는 도구인 Tableau이었다. 그 다음으로 많이 다루고 있는 주제는 프로그래밍 언어인 R

<표 11> 데이터 처리 및 분석: 프로그래밍 언어 교육 프로그램

교육 프로그램 내용		프로그램 수	프로그램 제공 기관
프로그래밍언어	R	10	보스턴대, 조지메이슨대, 존스홉킨스대, 노스캐롤라이나 주립대, 럿거스대, 텍사스대-알링턴, 콜로라도대, 노트르담대
	Python	6	보스턴대, 조지메이슨대, 콜로라도대, 예일대
	Perl	2	보스턴대
	C++	1	보스턴대
	CUDA	1	보스턴대
	GPU Programming	1	보스턴대
	Scalar	1	브라운대
합계		22	

〈표 12〉 데이터 시각화 및 표현 역량 관련 교육 프로그램

교육 프로그램 내용		프로그램 수	프로그램 제공 기관
시각화 도구	Tableau	10	보스턴대, 조지메이슨대, 조지아주립대, 노스캐롤라이나주립대, 텍사스대-알링턴, 예일대
	R	7	컬럼비아대, 듀크대, 노스캐롤라이나주립대, 럿거스대, 콜로라도대, 예일대
	Excel	1	브라운대
	PowerPoint	1	듀크대
	Adobe Photoshop	1	텍사스대-알링턴
	Adobe Illustrator	1	듀크대
	Maya	1	보스턴대
	CARTO	1	조지메이슨대
시각화 입문	4	브라운대, 라이스대, 텍사스대-알링턴, 버지니아커먼웰스대	
온라인 지도	3	듀크대, 노스캐롤라이나주립대, 콜로라도대	
스토리맵	2	듀크대, 아이오와주립대	
3D 시각화	1	앨버타대	
음성 시각화	1	노스캐롤라이나대	
합계	34		

을 활용한 시각화인 것으로 나타났다. 그 외에도 Maya나 CARTO와 같은 시각화 전문 소프트웨어부터 엑셀, 파워포인트, 포토샵 등 일반적인 소프트웨어를 활용한 시각화 교육 프로그램도 제공되고 있었다. 소프트웨어 이외에도 시각화에 대한 개괄적인 설명을 제공하는 입문 수업과 온라인 지도 또는 스토리맵 제작과 관련된 교육 프로그램들도 존재하였다.

6) 데이터 관리 및 조직

데이터 관리 및 조직과 관련된 교육 프로그램은 33개 기관에서 80개가 운영되고 있었다. 이를 통해 데이터 관리와 조직이 데이터 처리 및 분석 다음으로 많은 교육 프로그램이 제공되는 주제이면서 가장 많은 기관에서 교육하고 있는 역량임을 알 수 있다. 〈표 13〉에 따르면 데이터 관리에 대한 교육 프로그램은 25개 기관에서 50개가 운영되는 것으로 나타났는데 이중에서 데이터 관리의 개괄적이고 일반적인 내

용을 다루고 있는 교육 프로그램은 43개이었다. 그 외에도 데이터 관리를 분야별로 세분화하여 교육하는 기관들로 버지니아텍과 듀크대가 있었는데, 듀크대에서는 “Data Management 101”이라는 제목 하에 인문학, 사회과학, 자연과학으로 구분하여 워크숍을 제공하고 있었다. 교육대상별로는 미네소타대와 MIT에서 각각 대학원생과 박사후연구원을 대상으로 데이터 관리 교육을 운영하고 있었다. 데이터 유형별로는 정성적 데이터에 초점을 맞춘 데이터 관리 교육 프로그램이 있었다. 이 외에도 데이터 관리계획에 대한 교육 프로그램은 17개 기관에서 24개가 제공되고 있었으며 데이터 조직 관련 프로그램은 6개 기관에서 각각 하나씩 운영되고 있었다.

7) 데이터 품질 및 도큐멘테이션

데이터 품질 및 도큐멘테이션에 관련된 교육 프로그램은 4개 기관에서 각각 하나씩 제공되

〈표 13〉 데이터 관리 및 조직 역량 관련 교육 프로그램

교육 프로그램 내용		프로그램 수	프로그램 제공 기관
데이터 관리	일반	43	브라운대, 조지메이슨대, 존스홉킨스대, 루이지애나주립대, MIT, 노스웨스턴대, 라이스대, 애리조나대, 캘리포니아대-버클리, 캘리포니아대-데이비스, 캘리포니아대-리버사이드, 코네티컷대, 델라웨어대, 플로리다대, 일리노이대-어바나삼페인, 아이오와대, 메릴랜드대, 미네소타대, 미주리대, 노스캐롤라이나대, 버지니아텍, 오번대, 앨버타대, 토론토대
	분야별	4	듀크대, 버지니아텍
	교육대상별	2	MIT, 미네소타대
	데이터 유형별	1	캘리포니아대-버클리
데이터 관리계획		24	보스턴대, 브라운대, 조지메이슨대, 하버드대, 아이오와주립대, 존스홉킨스대, MIT, 노스캐롤라이나주립대, 테네시대, 델라웨어대, 플로리다대, 아이오와대, 메릴랜드대, 미네소타대, 노트르담대, 오레곤대, 버지니아텍
데이터 조직		6	보스턴대, 코넬대, 아이오와주립대, 존스홉킨스대, MIT, 아이오와대
합계		80	

는 것으로 나타나 그 수가 매우 적음을 알 수 있다. 존스홉킨스대와 일리노이대(어바나삼페인)에서 운영되는 교육 프로그램명에는 “data documentation”이라는 용어가 포함되어 있었으며, 다양한 유형의 데이터 도큐멘테이션을 소개하고 이를 작성하는 방법을 교육하고 있었다. MIT와 보스턴대에서 제공되는 교육 프로그램명에는 “version control”이 포함되어 있었는데 이는 데이터 품질 관리와 관련되며 데이터 세트의 버전을 추적·관리하는 내용을 다루고 있었다.

8) 메타데이터와 데이터 기술

메타데이터와 데이터 기술 관련 교육 프로그램의 경우에도 4개 기관에서 하나씩 제공되고 있어 제공빈도가 매우 적었다. 교육 내용으로 코넬대에서는 과학 분야 메타데이터(science metadata)에 대한 교육을 제공하고 있었고 존스홉킨스대에서는 효율적인 연구데이터 관리를 위해 메타

데이터를 포괄적인 범위에서 제시하는 방법에 대해 교육하고 있었다. MIT에서는 데이터의 검색과 해석 및 공유를 위해 메타데이터가 필수적임을 강조하고 이를 기술하는 방법에 대한 교육 프로그램을 운영하고 있었다. 캘리포니아대(버클리)에서는 데이터 기술의 중요성과 메타데이터의 개념 및 표준에 대해 교육하고 있었다.

9) 데이터 관리 및 공유 문화

데이터 관리 및 공유 문화와 관련해서는 6개 기관에서 10개 교육 프로그램이 제공되고 있었다. 주로 연구 재현성의 중요성과 이를 향상하기 위한 방법을 논의하고 있었으며 오픈 액세스, 연구 공개와 협력, 데이터 공유 시 고려사항에 대한 내용도 교육하고 있었다. 가장 많은 관련 프로그램을 운영하는 곳은 하버드대였는데 5개의 교육 프로그램을 운영하고 있었으며 그 중 3개 프로그램이 연구 재현성과 관련된 것이었다. 이들 프로그램에서는 의학 분야 연구자들을

위한 연구 재현성 관련 모범실무, 연구 재현성 향상을 위한 오픈 액세스 레포지터리 시스템인 protocol.io 및 기타 도구들에 대한 내용을 다루고 있었다. 나머지 2개 프로그램에서는 연구의 공개와 협력, 그리고 데이터 공유를 통한 이익과 우려 및 한계에 대한 내용을 교육하고 있었다. 그 외 5개 기관 중에서 보스턴대, 듀크대, 플로리다대에서는 연구재현성의 중요성과 이를 향상하기 위한 방법과 실무에 대한 교육 프로그램을 제공하고 있었다. 또한 콜로라도대에서는 오픈 액세스에 대한 기본 개념과 오픈 액세스를 통한 연구결과의 출판, 검색 및 공유에 대한 내용을, 코넬대에서는 데이터 공유 및 출판의 현황과 방법에 대한 내용을 다루고 있었다.

10) 윤리 및 데이터 인용

윤리 및 데이터 인용과 관련된 교육 프로그램은 8개 기관에서 11개가 제공되고 있었으며 그 내용은 <표 14>와 같다. 윤리와 관련된 것으로 데이터에서 개인정보를 제거하는 비식별화(de-identification)에 관련된 2개 프로그램이 있었으며 시각화와 관련된 윤리적 논의와 연구 윤리 및 데이터 보안에 관련된 교육 프로그램

들도 존재하였다. 데이터 인용과 관련해서는 데이터 인용에 대한 일반 사항을 제시하는 교육 프로그램과 데이터 식별과 인용을 지원하는 DOI(Digital Object Identifier) 및 ORCID(Open Researcher and Contributor ID)에 대한 프로그램들도 존재하였다.

11) 데이터 큐레이션 및 재이용

데이터 큐레이션 및 재이용 역량과 관련해서는 하버드대와 존스홉킨스대를 포함한 2개 기관에서 4개의 교육 프로그램이 운영되고 있었다. 하버드대에서는 데이터 큐레이션에 대한 전반적인 내용과 함께 하버드대에서 운영하고 있는 데이터 레포지터리인 Dataverse를 소개하는 프로그램을 운영하고 있었다. 존스홉킨스대에서는 연구데이터 재이용 및 소프트웨어의 연구 재현성과 재이용에 대한 3개의 교육 프로그램을 제공하고 있었다.

12) 데이터 보존

데이터 보존을 다루고 있는 교육 프로그램은 2개 기관에서 각각 하나씩 제공되고 있어 데이터 리터러시 세부 역량 중 가장 적은 수의 교육

<표 14> 윤리 및 데이터 인용 역량 관련 프로그램

교육 프로그램 내용		프로그램 수	프로그램 제공 기관
윤리	비식별화	2	존스홉킨스대
	시각화 관련 윤리	1	듀크대
	연구윤리	1	오레곤대
	데이터 보안	1	토론토대
데이터 인용	일반	2	보스턴대, 브라운대
	DOI	2	보스턴대, 플로리다대
	ORCID	2	캘리포니아대-데이비스, 플로리다대
합계		11	

프로그램이 운영되고 있는 것으로 나타났다. 교육 내용으로 존스홉킨스대에서는 데이터 백업 계획에 대한 안내를 제공하고 있었고 캘리포니아대(테이비스)에서는 데이터 복구에 대한 교육을 제공하고 있었다.

이와 더불어 집합교육을 제공하는 43개 기관 중 12개 기관에서는 워크숍에서 활용한 강의슬라이드 파일, 워크숍 활동자료 파일, 녹음된 강의슬라이드 및 강의동영상을 공개하고 있었다. 기관별로 공개하고 있는 교육 자료 유형과 빈도를 살펴보면 <표 15>와 같다. 1개 기관을 제외하고는 모두 강의 슬라이드 파일을 공개하고 있었으며 7개 기관에서는 워크숍 활동자료도 슬라이드와 함께 공개하고 있었다. 강의동영상을 제공하는 기관은 4곳이었으며 녹음된 강의 슬라이드를 제공하는 기관은 2곳이었다. 애리조나대의 경우 강의동영상은 해당 대학 ID를 가지고 있는 구성원들에 한해서만 공개하고 있었으며 나머지 대학들은 누구나 자료에 접근할 수 있도록 허용하고 있었다.

3.2.2 온라인 교육

집합 교육 없이 온라인 튜토리얼만을 제공하고 있는 기관은 본 연구의 조사대상인 51개 기관 중 8곳이었으며, 이 중에서 기관에서 자체적으로 개발한 온라인 튜토리얼을 제공하고 있는 기관은 5개인 것으로 나타났다. 나머지 3개 기관에서는 외부 기관에서 개발하여 공개하고 있는 온라인 튜토리얼 또는 교육 자료의 링크를 제공하고 있었다. 또한 집합 교육을 제공하는 43개 기관 중에서 외부 기관의 온라인 튜토리얼이나 교육 자료 링크도 함께 제공하고 있는 기관은 9곳이었다.

따라서 본 연구에서는 기관에서 자체 개발한 온라인 튜토리얼 5개를 대상으로 각 튜토리얼의 목차에서 가장 상위의 제목에 대해 데이터 리터러시 세부 역량을 기초로 코딩을 실시하였다. 또한 외부 온라인 튜토리얼 또는 교육 자료의 링크를 제공하는 12개 기관들을 대상으로 적어도 2개 이상의 기관에서 제시하고 있는 외부 온라인 교육 자료의 종류를 살펴보았다.

<표 15> 집합 교육 제공 기관에서 공개하고 있는 교육 자료 유형

	강의슬라이드 파일	녹음된 강의슬라이드	강의동영상	워크숍 활동자료 파일
듀크 대학	●		●	●
MIT	●			●
노스캐롤라이나주립대	●			●
노스웨스턴대	●			
테네시대(녹스빌)		●		
애리조나대	●		●	●
텔라웨어대	●		●	
플로리다대	●			
일리노이대(어바나삼페인)	●			●
아이오와대	●	●	●	●
미주리대	●			
예일대	●			●

자체 제작한 온라인 튜토리얼을 제공하는 5개 기관은 미시건주립대, 펜실베이니아주립대, 캘리포니아대(산타크루즈), 피츠버그대, 위스콘신대(밀워키)이었다. 각 기관의 튜토리얼 목차에서 수집된 가장 상위의 제목은 총 42개로 캘리포니아대(산타크루즈)의 목차 제목이 14개로 가장 많았고 피츠버그대에서 10개, 미시건주립대에서 8개, 나머지 두 대학에서 5개의 제목을 제공하고 있었다. 이들 제목을 대상으로 코딩을 수행한 결과 12개 데이터 리터러시 세부 역량 중 데이터 변환 및 상호운용성, 데이터 시각화 및 표현, 데이터 관리 및 공유 문화를 제외한 9개 리터러시 역량과 관련된 제목이 제시되고 있는 것으로 나타났다. <표 16>에 따르면 윤리 및 데이터 인용에 대한 제목이 10개로 가장 많았으며 데이터 보존 관련 제목이 9개, 데이터 관리 및 조직 관련 제목은 7개이었다.

이를 바탕으로 집합 교육에서는 거의 다루어지지 않은 데이터 보존 역량이 모든 조사대상 온라인 튜토리얼에서는 명시되고 있음을 확인할 수 있었다. 윤리 및 데이터 인용 역량의 경우 관련 제목 10개 중에서 캘리포니아대(산타크루즈)의 온라인 튜토리얼에서만 6개의 제목을 제공하고 있었으며 데이터 출판과 관련된 윤리적 문제, 데이터 라이선스, 데이터 인용과 DOI, ORCID, 대안지수(Altmetrics)에 대한 내용을 다루고 있었다. 윤리 및 데이터 인용 역량은 위스콘신대(밀워키)를 제외한 4개 기관의 튜토리얼에서 언급되고 있었으며 데이터 검색 및 수집 역시 같은 4개 기관 튜토리얼에서 다루어지고 있었다. 데이터 관리 및 조직 역량은 캘리포니아대(산타크루즈), 피츠버그대, 위스콘신대(밀워키) 등 3개 기관 튜토리얼에서 언급하고 있었다.

<표 16> 기관 자체 개발 온라인 튜토리얼 목차 제목 분석

역량	관련 제목 수	제목 예시(기관명)
데이터 검색 및 수집	4	Collect the data(미시건주립대)
		Finding Data(피츠버그대)
데이터베이스 및 데이터 포맷	1	Study the concept(미시건주립대)
데이터 처리 및 분석	2	Processing your data(미시건주립대)
		Analyzing your data(미시건주립대)
데이터 관리 및 조직	7	Characterizing Research Data and Research Data Management(피츠버그대)
		Organizing your data(위스콘신대-밀워키)
데이터 품질 및 도큐멘테이션	2	Documenting the data(펜실베이니아주립대)
메타데이터 및 데이터 기술	4	Data Description(캘리포니아대-산타크루즈)
		Workflow Documentation & Metadata Standards(캘리포니아대-산타크루즈)
윤리 및 데이터 인용	10	Policies for data sharing and access(펜실베이니아주립대)
		Citing and DOI's(캘리포니아대-산타크루즈)
데이터 보존	9	Long-term preservation and archiving of data(펜실베이니아주립대)
		Storing your data(위스콘신대-밀워키)
데이터 큐레이션 및 재이용	3	Repurposing your data(미시건주립대)
합계	42	

외부기관에서 개발된 온라인 튜토리얼의 링크를 제공하는 기관은 모두 12곳이었으며 이중 집합 교육 없이 외부 온라인 튜토리얼의 링크를 제공하는 기관은 유타대, 버지니아대, 워싱턴대 등 3곳, 집합 교육을 제공하고 있으면서 링크도 제공하는 기관은 9곳이었다. 이들 기관에서 링크로 제공되는 외부 온라인 튜토리얼은 모두 38개이었으며 이 중에서 2개 이상의 기관에서 제시하고 있는 온라인 튜토리얼은 6개인 것으로 나타났다. <표 17>에서와 같이 12개 기관 중 가장 많은 수인 8개 기관에서 공통적으로 제시하고 있는 외부 온라인 튜토리얼은 영국 에딘버러 대학 소속의 온라인 서비스 개발 센터인 Edina에서 구축한 무료 온라인 교육 모듈인 Research Data MANTRA 이었다. 그 다음으로 많은 6개 기관에서 공통적으로 제시한 온라인 튜토리얼은 DataONE Education Module이며 이를 개발한 DataONE(Data Observation Network for Earth)은 지구 관찰 데이터의 지속 가능한 관리를 위해 다양한 관련 기관들이 모여 구성한 협력체로서 미국국립과학재단(National Science Foundation)의 지원을 받고 있다. 그 다

음으로 3개 기관이 제시하고 있는 New England Collaborative Data Management Curriculum 의 경우 미국 매사추세츠대학(University of Massachusetts) 의학도서관에서 개발하였으며 보건과학, 자연과학 및 공학 분야 연구데이터 관리 모범실무를 교육하는 것을 목적으로 한다.

2개 기관들이 공통적으로 제시하는 온라인 교육 자료로 Research Data Management and Sharing MOOC는 노스캐롤라이나대학과 에딘버러 대학 EDINA의 협력 프로젝트에서 개발된 온라인 수업으로 Coursera를 통해 유료로 수강 가능하다. 미국지질조사국(U.S. Geological Survey)의 데이터 관리 교육 모듈은 과학 분야 데이터 관리에 초점을 맞추어 연구데이터 관리 전반에 대해 6개의 모듈로 설명하고 있다. 그 외 ICPSR Data Management and Curation는 데이터 관리 프로세스를 설명하고 있는 웹페이지이며 Data Management Snafu in 3 Short Acts의 경우 데이터 공유 및 재이용의 어려움을 3가지 에피소드로 구성한 유튜브 애니메이션 비디오이다.

<표 17> 2개 이상 기관에서 제시된 외부 온라인 튜토리얼 또는 교육 자료

외부 온라인 튜토리얼 또는 교육 자료	제공기관 수	제공기관
Research Data MANTRA	8	유타대, 버지니아대, 노스웨스턴대, 텍사스대(알링턴), 애리조나대, 미주리대, 오번대, 토론토대
DataONE Education Module	6	유타대, 버지니아대, 노스웨스턴대, 텍사스대(알링턴), 라이스대, 토론토대
New England Collaborative Data Management Curriculum	3	워싱턴대, 라이스대, 텍사스대(알링턴)
Research Data Management and Sharing MOOC	2	버지니아대, 토론토대
U.S. Geological Survey Data Management Training Modules	2	하버드대, 오번대
ICPSR Data Management and Curation	2	유타대, 토론토대
Data Management Snafu in 3 Short Acts	2	애리조나대, 하버드대

4. 요약 및 제언

본 연구에서는 북미의 대학도서관에서 제공되는 연구데이터 관리 교육 프로그램에서 다루어지고 있는 데이터 리터러시 세부 역량을 교육 제공 방식에 따라 구분하여 살펴보고 시사점을 제시하는 것을 목적으로 하였다. 분석 결과를 바탕으로 연구데이터 관리 교육 내용 및 제공 방식의 특징, the Carpentries와 같은 새로운 데이터 리터러시 교육 주체와의 협력 및 사서들을 대상으로 하는 소프트웨어 활용 및 데이터 관리 능력 교육에 대하여 논의하고자 한다.

첫째, 12개의 데이터 리터러시 세부 역량을 중심으로 집합 교육을 제공하는 43개 기관의 관련 교육 프로그램 제목 317개를 분석한 결과, 데이터 처리 및 분석 역량과 관련된 교육 프로그램의 개수가 24개 기관(55.8%)에서 140개(44.2%)로 가장 높은 비율을 차지하는 것으로 나타났다. 데이터 관리 및 조직 역량과 관련된 교육 프로그램의 경우 33개 기관(76.7%)에서 80개(25.2%)가 제공되는 것으로 나타나 가장 많은 기관들이 해당 역량과 관련된 교육 프로그램을 제공하고 있음을 알 수 있다. 또한 데이터 관리 및 조직 관련 교육 프로그램의 수도 두 번째로 많은 것으로 나타났다. 데이터 시각화 및 표현 역량과 관련된 교육 프로그램은 15개 기관(34.9%)에서 34개(10.7%)인 것으로 나타나 세 번째로 많은 수의 기관 및 프로그램에서 해당 역량을 다루고 있었다. 그러나 나머지 9개 데이터 리터러시 세부 역량을 중점적으로 다루고 있는 집합 교육 프로그램의 수는 매우 적은 것으로 나타났으며 이는 집합 교육 프로그램의 내용이 일부 데이터 리터러시 세부 역량에 편중되

어 있음을 보여준다. 교육 방식 역시 전체 51개 기관 중 43개 기관이 집합 교육을 제공하고 있었으므로 집합 교육이 일반적인 교육 제공 방식임을 알 수 있다. 그러나 집합 교육을 제공하는 기관 중에서도 12개 기관은 대다수 강의 슬라이드를 공개하고 있었으며 이 중 일부 기관에서는 강의동영상이나 녹음된 강의 슬라이드를 공개하고 있었다. 이는 대학 구성원뿐만 아니라 타 기관 또는 타 국가의 연구자들도 교육 프로그램에 접근할 수 있게 한다는 점에서 집합 교육의 한계를 보완하는 전략이 될 수 있다.

둘째, 전체 조사대상 기관 중 8개 기관은 집합 교육 없이 온라인 튜토리얼을 제공하고 있었으며 자체 개발한 온라인 튜토리얼을 제공하는 기관은 5곳이었다. 이들 튜토리얼의 목차 제목을 분석한 결과 데이터 보존, 윤리 및 데이터 인용, 데이터 관리 및 조직에 관한 제목들의 빈도가 높은 것으로 나타났다. 특히 데이터 보존의 경우 집합 교육에서는 관련 교육 프로그램의 수가 가장 적는데 비해 모든 온라인 튜토리얼에서 이를 명시하고 있어 교육 방식에 따라 제공되는 빈도에 큰 차이가 있음을 알 수 있다. 또한 대학도서관 교육 프로그램에서 일반적으로 강조되는 윤리와 인용에 대한 내용도 자체 개발된 온라인 튜토리얼에서는 많이 다루어지고 있었다. 이와 더불어 외부 기관에서 개발된 온라인 튜토리얼이나 교육 자료의 경우 집합 교육을 제공하는 9개 기관들도 참고 자료로써 링크를 제공하고 있었고 집합 교육을 제공하지 않는 3개 기관에서도 링크를 제시하고 있었다. 이와 같이 자체적으로 데이터 리터러시 세부 역량에 대한 교육을 제공하기 어려운 경우에는 무료로 제공되면서 연구데이터 관리 모범실무를 다루는 외부

온라인 튜토리얼을 활용하고 있음을 알 수 있다.

셋째, 분석 결과에도 나타났듯이 데이터 처리와 분석 역량은 대학도서관에서 전통적으로 교육해왔던 영역은 아니지만 이와 관련된 프로그램 수는 가장 많았다. 이를 운영하는 방법 중 하나로써 The Carpentries와 같은 외부 교육 기관을 활용하고 있는 대학들이 존재하였으며 본 연구에서도 적은 수이기는 하였으나 4개 기관이 이에 해당되었다. 이는 데이터 리터러시 교육과 관련된 새로운 현상이라고 볼 수 있으며 최근 제시된 학회 발표자료에서도 몇몇 관련 사례를 확인할 수 있었다. 예를 들어 미국 오클라호마 대학도서관의 경우 The Carpentries의 회원 기관으로 가입을 하여 워크숍을 주최함으로써 대학 내에서 소프트웨어 및 데이터 관련 교육을 제공하는 중심 기관 역할을 수행하고 있었다 (Clayton, 2018). 캘리포니아공과대학(CalTech) 도서관과 노르웨이의 오슬로 대학도서관에서도 연구데이터 관리 교육 프로그램의 일환으로 The Carpentries 워크숍을 활용하고 있었다 (Kvale & Stangeland, 2017; Wrublewski, Clement, & Morrell, 2017).

뿐만 아니라 이들 도서관에서는 사서들의 소프트웨어 및 데이터 능력 함양을 위한 계속교육을 Library Carpentry를 통해 제공하고 있었으며 본 연구에서도 1개 기관의 사례를 확인할 수 있었다. 특히 북미와 영국의 대학도서관에서는 연구데이터 관리가 새로운 실무 영역으로 부상하고 디지털 인문학 관련 과제를 많이 수행하게 되면서 사서들이 소프트웨어 및 프로그래밍 지식과 데이터 관리 및 분석 능력을 갖추어야 한다는 인식이 높아지고 있다(Baker, Moore, Priego, Alegre, Cope, Price, Stephens, Strien,

& Wilson, 2016; Cox, Verbaan, & Sen, 2012; Koltay, 2017; MacMillan, 2014). Library Carpentry에서는 이러한 지식의 습득을 지원함으로써 사서들의 커리어 개발을 도울 수 있으며 사서들의 전문지식 향상은 이용자들에게도 유용하다는 것을 강조하고 있다. 또한 사서들에게 유용한 관련 지식 영역으로 정규 표현식(regular expression), Unix Shell 환경에서의 명령줄 인터페이스(Shell/command-line interface), 버전 관리(version control), 데이터베이스, 데이터 정제, 스크립팅 언어(scripting languages)를 제시하였다(Cope & Baker, 2017). Library Carpentry는 아직 시작단계에 있으나 사서들이 연구데이터 관리 업무에 활용할 수 있는 기술적 지식과 능력을 습득하기 위한 계속교육 자원이 부족한 현실에서 의미 있는 시도라고 할 수 있다.

이와 같이 연구데이터 관리 교육 프로그램의 내용을 통해 북미 대학도서관 사서들에게 요구되는 지식 영역의 변화와 계속교육의 필요성, The Carpentries와 같은 IT 커뮤니티를 기반으로 하는 교육 주체와의 협력이 이루어지고 있는 현황을 살펴보았다. 국내에서도 연구기관이나 대학도서관에서 연구데이터 관리 서비스를 실행하기 위한 방안이 논의되고 있으며 이러한 서비스에서 교육 프로그램은 중요한 구성요소라고 할 수 있다. 이를 효과적으로 운영하기 위해서는 대학도서관 사서들이 전통적으로 교육하고 강조해왔던 영역뿐만 아니라 데이터 처리와 분석을 위한 소프트웨어 활용 및 프로그래밍 언어, 데이터 시각화와 표현 등 연구자들이 연구를 위해 필요로 하는 데이터 리터러시 세부 역량에 대한 이해와 이에 대한 지원이 요청된다고 할 수 있다. 새로운 지식 습득을 위한

사서들의 계속교육 기회와 교육 자원을 발굴·지원할 필요가 있으며 문헌정보학 정규 교과과정에서도 데이터 관리와 분석 역량 교육을 강화하기 위한 논의가 필요할 것이다.

5. 결론

북미의 대학도서관 연구데이터 관리 교육 프로그램은 대다수 집합 교육의 방식으로 이루어지고 있었으며 12개 데이터 리터러시 세부 역량 중 데이터 처리 및 분석, 데이터 관리 및 조직, 데이터 시각화 및 표현 역량에 중점을 두고 있었다. 따라서 이들 교육 프로그램에서 연구자들이 연구 결과를 도출하는데 필요한 역량과 연구비지원기관의 데이터 관리 요구사항과 관련된 교육에 초점을 맞추고 있음을 보여준다. 또한 나머지 9개 데이터 리터러시 세부 역량에 관련된 교육 프로그램의 수는 매우 적은 것으로 나

타나 교육 내용이 일부 역량에 편중되어 있음을 알 수 있다. 자체 개발한 온라인 튜토리얼을 제공하는 기관은 5곳이어서 집합 교육을 제공하는 기관 수에 비해 매우 적었다. 적은 수이기는 하지만 이들 온라인 튜토리얼의 목차 제목 분석을 통해 데이터 보존, 윤리 및 데이터 인용 등 대학도서관에서 전통적으로 강조하고 교육해 온 역량이 주로 언급되고 있다는 것을 확인할 수 있었다. 이러한 역량 역시 연구데이터 관리 교육 프로그램에서 강화되어야 할 내용이라고 할 수 있다. 이와 더불어 데이터 처리 및 분석, 시각화와 관련된 다양한 소프트웨어, 프로그래밍 언어 등 기술적 지식에 대한 사서들의 역량 강화와 계속교육 필요성이 북미 대학도서관 사서들 간 논의사항으로 부각되고 있음에 주목할 필요가 있다. 본 연구의 결과는 연구데이터 관리 서비스를 계획하는 국내 대학도서관 및 연구기관에서 교육 프로그램의 내용과 방식에 대한 기초자료로 활용될 수 있을 것이다.

참 고 문 헌

- 이재운 (2015). 데이터 사이언스와 데이터 리터러시. 한국정보관리학회 학술대회 논문집, 11-15.
- 한상우 (2017). 인문학 기반 데이터 리터러시 모형 설계에 관한 연구. 한국정보관리학회 학술대회 논문집, 45-71.
- Baker, J., Moore, C., Priego, E., Alegre, R., Cope, J., Price, L., ... & Wilson, G. (2016). Library Carpentry: software skills training for library professionals. *Liber Quarterly: The Journal of European Research Libraries*, 26(3), 141-162. <https://doi.org/10.18352/lq.10176>
- Brown, R. A., Wolski, M., & Richardson, J. (2015). Developing new skills for research support librarians. *The Australian library journal*, 64(3), 224-234.
- Bryant, R., Lavoie, B., & Malpas, C. (2017). Scoping the University RDM Service Bundle. *The Realities of Research Data Management, Part 2*. Dublin, OH: OCLC Research.

doi:10.25333/C3Z039.

- Carlson, J., Fosmire, M., Miller, C. C., & Nelson, M. S. (2011). Determining data information literacy needs: A study of students and research faculty. *portal: Libraries and the Academy*, 11(2), 629-657.
- Clayton, S. (2018). Software Carpentry in the Library [PowerPoint slides]. Retrieved from Re-think it 2018: Libraries for a New Age. <http://hdl.handle.net/2152/63437>
- Cope, J., & Baker, J. (2017). Library carpentry: software skills training for library professionals. *International Journal of Data Curation*, 12(1). <http://sro.sussex.ac.uk/69975/>
- Cox, A., Verbaan, E., & Sen, B. (2012). Upskilling liaison librarians for research data management. *Ariadne*, (70). <http://www.ariadne.ac.uk/issue70/cox-et-al>
- Faniel, I. M., & Connaway, L. S. (2018). Librarians' Perspectives on the Factors Influencing Research Data Management Programs. *College & Research Libraries*, 79(1), 100.
- Khan, H. R., Kim, J., & Chang, H.-C. (2018). Toward an Understanding of Data Literacy. Retrieved from <https://www.ideals.illinois.edu/handle/2142/100243>
- Koltay, T. (2016). Facing the Challenge of Data-Intensive Research: Research Data Services and Data Literacy in Academic Libraries. In *Innovation in Libraries and Information Services* (pp. 45-61). Emerald Group Publishing Limited.
- Koltay, T. (2017). Data literacy for researchers and data librarians. *Journal of Librarianship and Information Science*, 49(1), 3-14.
- Kvale, L., & Stangeland, E. (2017). Skills for research data management-developing RDM courses at the university of Oslo. *Proceedings of the Association for Information Science and Technology*, 54(1), 728-730.
- MacMillan, D. (2014). Data sharing and discovery: What librarians need to know. *The Journal of Academic Librarianship*, 40(5), 541-549.
- Nelson, M. S. (2015). Where do we go from here? Furthering developing the data information literacy competencies. In Carlson, J., & Johnston, L. R. (Eds.). *Data information literacy: Librarians, data, and the education of a new generation of researchers* (pp.231-245). Purdue University Press.
- Prado, J. C., & Marzal, M. Á. (2013). Incorporating data literacy into information literacy programs: Core competencies and contents. *Libri*, 63(2), 123-134.
- Qin, J., & D'Ignazio, J. (2010). Lessons learned from a two-year experience in science data literacy education. Paper presented at 31st Annual IATUL Conference, June 22, 2010. <http://docs.lib.purdue.edu/iatul2010/conf/day2/5>

- Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., ... & Wuetherick, B. (2015). Strategies and best practices for data literacy education: Knowledge synthesis report. Dalhousie University.
- Schneider, R. (2013). Research data literacy. In Kurbanoglu, S. et al. (Ed.), *Worldwide Commonalities and Challenges in Information Literacy Research and Practice*, Springer International, Cham, pp. 134-140.
- Shorish, Y. (2015). Data information literacy and undergraduates: A critical competency. *College & Undergraduate Libraries*, 22(1), 97-106.
- Teal, T.K., Cranston, K.A., Lapp, H., White, E., Wilson, G., Ram, K., & Pawlik, A. (2015). Data carpentry: Workshops to increase data literacy for researchers. *International Journal of Digital Curation*, 10(1), 135-43. doi:10.2218/ijdc.v10i1.351
- Tenopir, C., Talja, S., Horstmann, W., Late, E., Hughes, D., Pollock, D., ... & Allard, S. (2017). Research data services in European academic research libraries. *Liber Quarterly*, 27(1), 23-44. <http://doi.org/10.18352/lq.10180>
- Wrublewski, D., Clement, G., & Morrell, T. (2017). Data carpentry in the Caltech libraries. Paper presented in 254th American Chemical Society National Meeting & Exposition, August 20-24, 2017, Washington, DC.

[웹사이트]

- Center for Open Science. n.d. Manage and share your research with OSF. Retrieved from <https://cos.io/our-products/osf/>
- Data Carpentry. n.d. Data Carpentry Vision and Mission. Retrieved from <https://datacarpentry.org/mission/>
- Library Carpentry. n.d. About Library Carpentry. Retrieved from <https://librarycarpentry.org/>.
- The Carpentries. n.d. About us. Retrieved from <https://carpentries.org/about/>

• 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean)

- Han, Sang Woo (2017). A research on the design of data literacy model based on digital humanities. *Proceedings of the Korean Society for Information Management*, 45-71.
- Lee, Jae Yun (2015). Data science and data literacy. *Proceedings of the Korean Society for Information Management*, 11-15.

[부록 1] 51개 연구데이터 관리 교육프로그램 제공 기관

	카네기분류 R1 and/or ARL	국가	교육제공 방식	프로그램 수	교육제공 시기	비고
Boston University (보스턴대)	R1 & ARL	미국	집합교육	32	2016.9- 2017.11	
Brown University (브라운대)	R1 & ARL	미국	집합교육	10	N/A	
Columbia University (콜럼비아대)	R1 & ARL	미국	집합교육	5	2018.9- 2018.11	
Cornell University (코넬대)	R1 & ARL	미국	집합교육	6	2018.9- 2018.11	
Duke University (듀크대)	R1 & ARL	미국	집합교육	23	2018.9- 2018.11	
GeorgeMasonUniversity (조지메이슨대)	R1	미국	집합교육	31	N/A	
Georgia State University (조지아주립대)	R1	미국	집합교육	17	2018.9- 2018.11	
Harvard University (하버드대)	R1 & ARL	미국	집합교육	9	2018.9- 2018.11	외부기관 온라인 튜토리얼 링크 제공
Iowa State University (아이오와주립대)	R1 & ARL	미국	집합교육	4	2018.9- 2018.11	
Johns Hopkins University (존스홉킨스대)	R1 & ARL	미국	집합교육	22	2018.9- 2018.11	외부기관 온라인 튜토리얼 링크 제공
Louisiana State University (루이지애나주립대)	R1 & ARL	미국	집합교육	1	2018.9- 2018.11	
MIT	R1 & ARL	미국	집합교육	9	N/A	
North Carolina State University at Raleigh (노스캐롤라이나주립대)	R1 & ARL	미국	집합교육	19	2018.9- 2018.11	
Northwestern University (노스웨스턴대)	R1 & ARL	미국	집합교육	4	2014- 2016	외부기관 온라인 튜토리얼 링크 제공
Rice University (라이스대)	R1 & ARL	미국	집합교육	5	2018.6- 2018.8	외부기관 온라인 튜토리얼 링크 제공
Rutgers University-New Brunswick (럿거스대)	R1 & ARL	미국	집합교육	3	N/A	
Stanford University (스탠포드대)	R1	미국	집합교육	3	2018.9- 2018.11	
The University of Tennessee-Knoxville (테네시대)	R1 & ARL	미국	집합교육	2	2018.9- 2018.11	
The University of Texas at Arlington (텍사스대-알링턴)	R1	미국	집합교육	10	2018.9- 2018.11	외부기관 온라인 튜토리얼 링크 제공
University of Arizona (애리조나대)	R1 & ARL	미국	집합교육	3	2017.9- 2018.3	외부기관 온라인 튜토리얼 링크 제공
University of California-Berkeley (캘리포니아대-버클리)	R1 & ARL	미국	집합교육	2	N/A	
University of California-Davis (캘리포니아대-데이비스)	R1 & ARL	미국	집합교육	15	2016.5- 2018.10	
University of California-Riverside (캘리포니아대-리버사이드)	R1 & ARL	미국	집합교육	6	2018.9- 2018.11	
University of California-San Diego (캘리포니아대-샌디에고)	R1 & ARL	미국	집합교육	1	2018.9- 2018.11	
University of Colorado Boulder (콜로라도대)	R1 & ARL	미국	집합교육	17	2018.9- 2018.11	
University of Connecticut (코네티컷대)	R1 & ARL	미국	집합교육	2	2018.9- 2018.11	

	카네기분류 R1 and/or ARL	국가	교육제공 방식	프로그램 수	교육제공 시기	비고
University of Delaware (델라웨어대)	R1 & ARL	미국	집합교육	2	2015.10-2016.11	
University of Florida (플로리다대)	R1 & ARL	미국	집합교육	6	2016.9-2018.3	
University of Illinois at Urbana-Champaign (일리노이대-어바나샴페인)	R1 & ARL	미국	집합교육	3	2018.9-2018.11	
University of Iowa (아이오와대)	R1 & ARL	미국	집합교육	6	2015.11-2017.5	
University of Maryland-College Park (메릴랜드대)	R1 & ARL	미국	집합교육	2	2018.3	
University of Minnesota-Twin Cities (미네소타대)	R1 & ARL	미국	집합교육	3	N/A	
University of Missouri-Columbia (미주리대)	R1 & ARL	미국	집합교육	1	2016.11	
University of North Carolina at Chapel Hill (노스캐롤라이나대)	R1 & ARL	미국	집합교육	1	N/A	
University of Notre Dame (노트르담대)	R1 & ARL	미국	집합교육	3	2018.9-2018.11	
University of Oregon (오레곤대)	R1 & ARL	미국	집합교육	2	N/A	외부기관 온라인 튜토리얼 링크 제공
University of Wisconsin-Madison (위스콘신대-매디슨)	R1 & ARL	미국	집합교육	2	2018.9-2018.11	
Virginia Commonwealth University (버지니아 커먼웰스대)	R1 & ARL	미국	집합교육	1	2018.9-2018.11	
Virginia Polytechnic Institute and State University (버지니아텍)	R1 & ARL	미국	집합교육	7	2015.2-2018.2	
Yale University (예일대)	R1 & ARL	미국	집합교육	8	2018.9-2018.11	
Auburn University (오번대)	ARL	미국	집합교육	1	N/A	외부기관 온라인 튜토리얼 링크 제공
University of Alberta (앨버타대)	ARL	캐나다	집합교육	4	2018.5	
University of Toronto (토론토대)	ARL	캐나다	집합교육	4	N/A	외부기관 온라인 튜토리얼 링크 제공
Michigan State University (미시건주립대)	R1 & ARL	미국	자체개발 온라인 튜토리얼			
Pennsylvania State University (펜실베이니아주립대)	R1 & ARL	미국	자체개발 온라인 튜토리얼			
University of California-Santa Cruz (캘리포니아대-산타크루즈)	R1	미국	자체개발 온라인 튜토리얼			
University of Pittsburgh-Pittsburgh Campus (피츠버그대)	R1 & ARL	미국	자체개발 온라인 튜토리얼			
University of Wisconsin-Milwaukee (위스콘신대-밀워키)	R1	미국	자체개발 온라인 튜토리얼			
University of Utah (유타대)	R1 & ARL	미국	외부기관 온라인 튜토리얼 링크			
University of Virginia-Main Campus (버지니아대)	R1 & ARL	미국	외부기관 온라인 튜토리얼 링크			
University of Washington-Seattle Campus (워싱턴대)	R1 & ARL	미국	외부기관 온라인 튜토리얼 링크			