

# 디지털 큐레이션 연구동향 분석과 과제: 문헌정보학 분야를 중심으로

## An Analytical Study on Research Trends of Digital Curation: Focused on Library and Information Science

김판준 (Pan Jun Kim)\*

### 초 록

디지털 큐레이션은 디지털 시대의 새로운 정보관리 및 서비스라고 할 수 있다. 디지털 환경에서 정보자원에 대한 체계적 관리 및 서비스를 추구한다는 측면에서 디지털 큐레이션은 문헌정보학의 핵심 영역에 해당한다. 본 연구는 디지털 큐레이션 관련 문헌들을 검토하여 주요 연구 영역을 설정하고, 대표적인 문헌정보학 분야 학술 데이터베이스(LISTA)에서 검색한 학술 논문들을 대상으로 연구동향을 분석하였다. 또한 이러한 연구동향 분석의 결과에 기초하여 문헌정보학 분야의 디지털 큐레이션 관련 향후 연구과제를 제시하였다.

### ABSTRACT

Digital curation can be said that the new information management and services in the digital age. In terms of exploring the systematic management and services for the information resources in the digital environment, Digital curation is one the core areas of library and information science. This study was set up a research areas as a result of reviewing the related literatures, and analyzed the research trends to the scholarly articles retrieved from a representative databases in the areas of Library and Information Science (LISTA). Also, I suggested future research agendas for digital curation in the areas of library and information science based on the results of the this analysis.

키워드: 디지털 큐레이션, 정보관리 및 서비스, 정보서비스, 문화유산, 기관, 지적구조 분석, 연구동향, 연구과제  
digital curation, information management and services, information services, cultural property, institutions, intellectual structure analysis, research trends, research agendas

---

\* 신라대학교 문헌정보학과 조교수(pjkim@silla.ac.kr)

■ 논문접수일자: 2015년 2월 26일 ■ 최초심사일자: 2015년 2월 26일 ■ 게재확정일자: 2015년 3월 17일  
■ 정보관리학회지, 32(1), 265-295, 2015. [http://dx.doi.org/10.3743/KOSIM.2015.32.1.265]

## 1. 서론

### 1.1 연구의 필요성과 목적

디지털 시대의 도래와 함께 다양한 분야에서 디지털 정보의 생산과 축적이 일상화되고 있는 가운데 이러한 정보자원의 체계적인 관리와 활용에 대한 요구가 증가하고 있다. 그러나 다양한 형식과 엄청난 규모로 생산되고 있는 디지털 정보의 체계적 수집과 관리 및 서비스에 대한 논의는 상대적으로 부족한 실정이다. 이에 따라 장기간의 접근가능성을 보장하기 위한 디지털 정보의 관리 및 보존에 초점을 둔 디지털 큐레이션에 대한 관심이 커지고 있다(Abbott, 2008). 디지털 큐레이션은 반복적인 워크숍과 아젠다 설정 과정을 통해 새로운 학문분야로 부상하였으며, 특히 영국에서는 디지털 정보의 장기적 관리에 대한 전략적 초점이 점진적으로 수동적 보존에서 동적 큐레이션으로 이동하였다. 정의와 안정화(consolidation)의 시기 이후, 디지털 큐레이션은 최근 국제적 전문직의 성장, 연구 아젠다와 실무적 도구 및 협력 프로젝트의 개발, 고등교육 수준으로 훈련된 인력의 양성이 활발히 진행되고 있다(Higgins, 2011). 이처럼 디지털 큐레이션은 초기의 이론적 논의를 거쳐 디지털 정보의 체계적인 보존 및 관리에 기초한 장기적 접근 및 재사용을 보장하는 디지털 시대의 새로운 정보서비스로 발전하고 있다.

디지털 환경에서 정보자원에 대한 체계적 관리 및 서비스를 추구한다는 측면에서 본질적으로 디지털 큐레이션은 문헌정보학의 핵심 영역에 해당한다. 다만, 전통적인 문헌정보학 분야

의 정보관리 및 서비스가 아날로그 환경에서 도서관 및 인쇄자료에 초점을 두고 있는 것과 달리, 디지털 큐레이션은 디지털 환경을 전제로 보다 확장된 의미에서 다양한 문화유산기관 및 디지털 자원을 대상으로 하고 있다는 점에서 차이가 있다. 즉, 디지털 큐레이션은 도서관뿐 아니라 리포지토리, 박물관, 기록관 등 여러 문화유산기관들이 다양한 유형과 형식의 디지털 정보에 대하여 생애주기 동안 접근 및 재사용을 보장하기 위한 일련의 활동이라 할 수 있다. 여기서 정보자원의 형식이 단순히 아날로그에서 디지털로 전환되는 차원을 넘어 정보관리 및 서비스의 주체는 물론 그 범위와 내용이 본질적으로 변화되고 있음에 주목해야 한다. 따라서 이 시점에 문헌정보학 분야에서 지금까지 디지털 큐레이션이라는 새로운 영역을 어떻게 접근 및 탐구하고 있는지를 살펴보고, 이를 통해 정보관리 및 서비스의 전통적인 주체로서 효율적으로 대응하기 위한 미래 전략을 마련할 필요가 있다. 따라서 본 연구는 문헌정보학 분야를 중심으로 디지털 큐레이션 관련 연구동향을 분석하고, 그 결과에 기초하여 차세대 디지털 정보관리 및 서비스의 주체로서 위상을 확립하기 위한 향후의 연구과제를 제시해 보고자 한다.

### 1.2 연구 방법 및 제한점

본 연구의 연구방법은 크게 세 단계로 구분할 수 있다. 첫째, 주요 서지, 논문, 단행본 등 디지털 큐레이션 관련 문헌들의 내용을 검토하여 주요 연구 영역을 설정하였다. 둘째, 문헌정보학 분야의 대표적인 학술 데이터베이스 중 하나인 LISTA에서 검색한 디지털 큐레이

션을 다룬 학술지 논문들을 대상으로 연구동향을 분석하였다. 셋째, 앞서 설정한 주요 연구 영역별로 실제 연구동향을 분석한 결과를 토대로 향후 문헌정보학 분야에서 중점적으로 다루어야 할 연구과제를 제시하였다. 여기서 문헌정보학 분야 디지털 큐레이션 관련 연구동향은 기본 분석과 지적구조 분석을 통하여 파악하였다. 특히 지적구조 분석으로는 통제어휘와 비통제어휘 각각의 장점을 살리고 단점은 보완할 수 있는 디스크립터 프로파일링(김관준, 이재윤, 2007) 기법에 기초한 시기별·영역별 분석과 생산성과 성장성을 고려한 종합 분석을 수행하였다.

본 연구의 제한점으로는 첫째, 전체 학문 분야를 대상으로 하지 않고 문헌정보학 분야를 중심으로 국외의 연구동향을 살펴보았다. 둘째, 분석 대상이 되는 문헌집합을 학술지 논문으로 제한하여 학위논문, 단행본 등은 포함하지 않았다. 셋째, 학술지 논문에 부여된 디스크립터를 중심으로 분석하였고, 다른 관점에서 연구동향을 파악할 수 있는 연구자 또는 연구집단에 대한 분석은 수행하지 않았다.

### 1.3 용어 정의

디지털 큐레이션은 비교적 최근에 부상하고 있는 분야로서 문헌정보학은 물론 컴퓨터학, 경영학 등 여러 학문분야에서 다양한 관점으로 연구되고 있다. 이에 따라 학문분야들 간에 합의되어 통용되는 용어가 많지 않고 그 정의나 의미가 모호한 경우가 많은 문제가 있다. 따라서 본 연구에서 사용된 주요 용어들에 대한 정의를 제시하였다.

- **디지털 정보(자원):** 디지털 형식으로 생산 및 유통되는 유형·무형의 모든 정보(자원)로서, 정보관리 및 서비스의 대상이 될 수 있는 인식 가능한 디지털 정보 객체를 말한다.
- **디지털 정보서비스:** 디지털 정보(자원)의 체계적인 관리에 기초한 새로운 정보서비스로서 디지털 정보관리 및 서비스와 동일한 의미이다. 초기에는 수집, 정리, 보존, 저작권 등 정보관리에 중점을 두지만 시간의 경과에 따라 정보관리와 병행하는 접근 및 재사용, 교육 및 훈련, 연구(계획, 출판, 인용 등), 기획 및 분석 등의 서비스 중심으로 발전한다. 이러한 맥락에서 디지털 큐레이션은 새롭게 부상하고 있는 차세대 디지털 정보서비스라고 할 수 있다.
- **디지털 문화유산:** 문화 발전을 위해 다음 세대에 물려줄 만한 가치를 지닌 과학, 기술, 관습, 규범 따위의 민족 사회 또는 인류 사회의 문화적 소산이 되는 디지털 정보(자원)을 말한다.
- **디지털 문화유산기관:** 디지털 문화유산의 관리 및 서비스를 주도적으로 수행하는 기관으로, 전통적인 문화유산기관인 도서관, 기록관, 박물관에 더하여 디지털 환경에서 새롭게 출현한 리포지토리 등을 포함한다.
- **디지털 큐레이션:** 생애주기 동안 디지털 정보(자원)의 체계적인 관리를 통하여 (영속적인) 접근 및 재사용을 보장하는 디지털 정보서비스. 여기서 디지털 정보(자원)의 관리 및 서비스는 선정 및 수집, 보존 및 아카이빙, 접근 및 재사용, 권한 관리 및 보안, 메타데이터, 파일 포맷 및 인코딩, 교

육 및 훈련, 홍보, 연구(계획, 출판, 인용 등), 부가서비스(기획 및 분석 등) 등을 포함한다. 또한, 디지털 큐레이션의 대상이 되는 정보(자원)는 기존의 도서관자료, 기록물, 박물관과 함께 데이터를 포괄하는 디지털 문화유산으로 확장되고 있다.

- **디지털 보존 또는 디지털 아카이빙:** 지속적으로 보존할 가치를 가진 디지털 객체를 장기간 관리하여 이후의 이용을 보장하는 활동. 디지털 보존이 구체적인 보존 기술이나 전략을 지칭하는 경우가 많은 데 비하여, 아카이빙은 보존 행위의 주체가 되는 아카이브의 구축 및 운영까지에 이르는 더 넓은 의미로 사용되는 경향이 있다(한국기록학회, 2008).
- **인프라(Infrastructure):** 디지털 큐레이션의 토대가 되는 정보기술 환경으로서 e-Science, cyberinfrastructure, e-research 등을 통칭한다. 1999년 영국의 John Taylor 경(Director General of the UK Research Councils)이 주요 과학 분야에서의 글로벌 협력과 전자 형식으로 최대한 접근하도록 지원할 차세대 인프라로 e-Science를 언급하였고 이후, 미국에서는 '사이버인프라(cyberinfrastructure)'라는 용어를 데이터 수집, 데이터 저장, 데이터 관리, 데이터 통합, 데이터 마이닝, 데이터 시각화, 인터넷을 통한 컴퓨팅과 정보처리 서비스를 지원하는 새로운 연구 환경으로 사용하였다. 또한 'e-research'는 호주와 다른 국가들에서 인문학과 사회과학을 포함하는 것으로 e-Science와 사이버인프라의 범위를 확장하여 기존 및 새로운 형식의 연구를 지원하

는 정보기술의 이용을 의미한다(Simons & Richardson, 2013, pp. 20-21).

- **스몰 데이터:** 개인이나 중·소규모의 집단에 의해 수행되는 연구에서 생산되는 빅데이터에 비해 소규모(또는 중규모)이며 개인 컴퓨터에서 이용 및 조작이 가능한 데이터 또는 데이터셋을 말한다(Carlson, 2006; Gold, 2007; Purdue University Libraries, 2007; Choudhury, 2008; Steinhart, Dietrich, & Green, 2009).

## 2. 이론적 배경

### 2.1 디지털 큐레이션의 개념 및 의의

디지털 큐레이션은 디지털 시대의 모든 문화유산기관들이 제공해야 할 새로운 디지털 정보서비스이다. 이러한 디지털 정보서비스는 사전에 체계적인 정보관리가 전제된다는 점에서 정보 저장 및 검색(Information Storage and Retrieval)이 정보검색(Information Retrieval)으로 전환된 것과 동일한 맥락에서 정보관리 및 서비스에서 정보서비스로 성장하는 과정에 있다. 즉, 정보검색이 구조화된 저장을 전제로 하는 것과 마찬가지로 디지털 정보의 체계적인 관리가 선행되어야 효율적인 서비스가 가능하다는 점에서, 디지털 정보자원의 관리 및 서비스가 디지털 정보서비스로 전환 및 발전하는 과정에 있는 것이다.

디지털 큐레이션은 2001년 10월 19일 런던에서 Digital Preservation Coalition and the British National Space Centre 주최로 개최한

'Digital Curation: digital archives, libraries and e-science seminar'에서 처음으로 사용되었다. 여기서 디지털 큐레이션은 디지털 자료를 수집, 유지, 영구보존, 제공하는 아카이빙 활동을 의미하며, 데이터가 학술적인 관심 대상이 되는 생명주기동안 이러한 데이터를 적극적으로 관리하고 평가하는 것으로, 이는 데이터의 재현과 재사용을 위해 반드시 필요한 영역이라고 하였다(DCC, 2013). 따라서 데이터 큐레이션도 디지털 큐레이션의 하위 범주라 할 수 있으며, 그 개념은 데이터 사용을 관리하는 활동으로 생산시점에서부터 데이터의 정보자원으로서의 가치를 발견하고 미래에도 재사용할 수 있도록 하는 활동이다(안영희, 박옥화, 2010). 이와 같이 데이터 큐레이션과 관련해서는 디지털 아카이빙, 디지털 보존과 함께 디지털 자원 관리로써 디지털 큐레이션의 하위 범주로 보는 견해가 일반적이다(신영란, 정연경, 2012).

아날로그 환경의 정보서비스와는 달리 컴퓨터 처리 및 네트워크 서비스를 전제로 하는 디지털 환경의 새로운 정보서비스로서 디지털 큐레이션은 크게 세 가지 차원에서 그 의미가 전환 및 확장되고 있다. 첫째, 정보서비스의 대상이 되는 정보자원 측면에서 디지털 큐레이션은 기존의 도서관자료, 기록물, 박물관에 더하여 데이터를 포함하는 디지털 문화유산으로 확장되었다. 특히, 디지털 융합 현상으로 인하여 기관 및 컬렉션 측면에서의 도서관자료, 기록물, 박물관 등의 구분은 그 의미가 축소되고 있는 반면(이소연, 2012), 새로운 정보자원으로서 데이터의 중요성이 커지고 있다. 심지어는 디지털 큐레이션과 데이터 큐레이션이 여러 문헌에서 상

호교환이 가능한 동등한 의미로 사용되고 있다(Gold, 2010). 예를 들면, National Science Board(2005)는 데이터를 포맷에 상관없이 디지털 형식으로 저장된 어떤 정보라 하여 양자를 동일한 것으로 간주한다. 그러나 본 연구에서 디지털 큐레이션은 디지털 형식으로 생산된 모든 유형의 정보에 대한 큐레이션을 말하며, 데이터 큐레이션은 이의 하위 개념으로 디지털 정보 중에서도 데이터에 대한 큐레이션을 말한다. 이처럼 디지털 큐레이션이 데이터 큐레이션을 포함하는 것으로 정의하는 것은 디지털 큐레이션 커뮤니티가 정적인 디지털 객체를 넘어 구조, 휘발성, 스케일을 갖는 자원들로 큐레이션을 확장하는 것이다(Gold, 2010). 따라서 디지털 큐레이션은 원래 디지털 형식으로 생산(born-digital)되거나 아날로그 자료가 디지털화(digitization)된 정보자원을 포함하여 모든 유형의 디지털 객체를 대상으로 한다. 특히, 인쇄정보 중심의 아날로그 환경에서는 정보서비스의 대상으로서 뚜렷한 지위를 갖지 못했던 데이터가 디지털 환경에서는 독립적인 정보 유형으로서 상당한 의미를 갖는다.

둘째, 정보서비스의 범위 및 내용 측면에서 디지털 큐레이션은 생애주기 동안 디지털 정보(자원)의 체계적인 관리를 통하여 (영속적인) 접근 및 재사용을 보장하는 것으로 확장되고 있다. 여기서 관리 및 서비스의 내용으로는 디지털 문화유산의 선정 및 수집, 보존 및 아카이빙, 접근 및 재사용, 권한 관리 및 보안, 메타데이터, 파일 포맷 및 인코딩, 교육 및 훈련, 홍보, 연구(계획, 출판, 인용 등), 부가서비스(기획 및 분석 등)를 포함한다. 즉 디지털 큐레이션은 "안전하고 온전하게 유지하고 재사용을 위한

접근이 가능하도록 디지털 정보의 보존, 유지, 아카이빙, 보관을 포함하는 일련의 활동”으로서, 시작(생산)에서부터 새로운 지식 상품으로의 변환을 위한 재사용에 이르기까지 데이터의 전체 생애주기를 포괄한다(Jahnke et al., 2012). 디지털 아카이빙과 디지털 보존은 모두 큐레이션 관련 활동의 일부로 볼 수 있으며, 디지털 큐레이션은 데이터 큐레이션, 디지털 아카이빙, 디지털 보존을 포괄하는 디지털 정보 자원의 관리 및 서비스라 할 수 있다. 따라서 디지털 큐레이션은 디지털 자원의 보존 및 관리뿐만 아니라 생산 단계부터 포괄적인 활용까지의 모든 과정을 포함한다는 점에서 디지털 아카이빙보다 넓은 범주의 활동이다. 특히, 데이터를 포함하는 디지털 자원에 대한 적용에서 큐레이션은 재사용을 용이하게 하고 책임과 통합을 촉진하는 방식으로 맥락과 연결을 제공함으로써 가치를 추가하는 관리(stewardship)를 포함하는 현재의 재사용 증진과 아카이빙 책임을 넘어서고 있다(Rusbridge et al., 2005, p. 2).

셋째, 정보서비스를 제공하는 주체(기관) 측면에서 디지털 큐레이션은 전통적인 문화유산 기관인 도서관, 기록관, 박물관에 더하여 리포지토리를 포함하는 것으로 확장 또는 통합되고 있다. 전자의 3개 기관은 대표적인 문화유산기관으로서 서로 다른 유형의 정보자원을 수집하여 구축한 컬렉션을 기반으로 정보서비스를 제공한다는 공통점을 갖는다. 그러나 디지털 환경이 보편화되면서 이전의 아날로그 환경과 달리 간행물이나 기록물, 3차원 객체 등으로 정보자원을 구분하지 않고 정보관리 및 서비스가 이루어지는 디지털 융합 현상이 발생하고 있다. 실

제로 1998년 도서관, 기록관, 박물관 등의 문화유산기관들 간의 디지털 융합이 최초로 예고된 이후, 2004년에 캐나다 국립도서관과 국립기록관이 통합되고, 비슷한 시기에 영국의 국립기록관과 역사위원회가 통합되었다. 최근에는 한국 원자력연구원이 도서관, 기록관, 박물관을 통합한 라키비움을 설립하여 서비스를 제공하고 있다(이소연, 2012). 따라서 디지털 시대의 대표적인 문화유산기관으로서 도서관, 기록관, 박물관, 리포지토리는 디지털 큐레이션의 핵심 주체(기관)가 되고 있다.

## 2.2 문헌정보학과 디지털 큐레이션

문헌정보학에서 정보관리와 서비스는 가장 핵심적인 영역이라 할 수 있으며, 전통적으로 도서관은 중립적이고 유능한 정보 브로커로서 한결같은 신뢰성을 확보하고 있다(McCue and Corson-Rikert, 2007). 그러나 1990년대 중반부터 인터넷을 매개로 시작된 디지털 환경으로의 전환 과정에서 문헌정보학은 정보관리 및 서비스 측면에서 타 학문 또는 산업 분야의 강력한 도전에 직면해 있으며, 대표적인 정보제공기관으로서 도서관의 위상도 나날이 위축되는 상황에 있다. 디지털 형식으로 생산되는 정보의 양과 다양성이 크게 증대하면서 정보를 입수할 수 있는 매체와 경로가 다양해지고, 보다 편리한 방식으로 직·간접적인 정보서비스를 제공하는 조직들이 늘어나면서 이러한 현상은 더욱 가속화되고 있다. 이용자들은 항상 정보의 유형과 형식에 상관없이 적합한 정보를 적시에 찾고 원하는 목적과 방식으로 이용하기를 원한다. 즉, 이용자들은 정보관리 및 서비스를 제공하는 주

체가 누구이든 자신에게 가장 편리한 방식으로 적합한 정보를 적시에 입수하여 원하는 목적으로 활용할 수 있는가에 관심이 있을 뿐이다. 따라서 문헌정보학은 이러한 시대적 조류와 이용자들의 요구를 반영한 정보관리 및 서비스를 제시할 수 있어야 하며, 도서관은 이를 실질적으로 구현할 수 있는 시스템과 서비스를 제공하여야 한다. 이를 위해 문헌정보학은 새로운 환경과 이용자 요구에 적절한 이론을 정립하고, 이에 기초하여 차세대 정보관리 및 서비스의 방향 및 내용을 제시하여야 한다. 또한, 도서관은 고대 도서관들이 그러하였듯이 도서관, 박물관, 기록관이 통합된 문화유산기관으로서 정보관리 및 서비스를 제공할 수 있어야 하며, 디지털 환경에서 새롭게 출현한 정보제공기관 또는 문화유산기관들도 수용하여야 한다. 기존의 아날로그 환경과는 달리 모든 유형과 형식의 정보들이 디지털 형식으로 융합되는 디지털 환경에서, 도서관은 디지털 문화유산의 통합 관리를 기반으로 하는 새로운 정보서비스의 핵심 주체로 거듭나야 한다. 1990년대 초반부터 2000년대 초반에 이르기까지 미국은 국가적인 디지털 도서관 사업을 대대적으로 추진하는 가운데 기술 및 컬렉션(콘텐츠)에서 이용자와 서비스 중심으로 전환하는 과정을 거쳤다. 이러한 맥락에서 2000년대 중반 이후 NSF가 추진한 디지털 큐레이션 사업에서 가장 중점을 둔 것은 서비스의 구축으로서 데이터의 수집과 보존은 중간 목표이며, 보다 근본적인 목적은 재사용과 학문분야간 사용을 위한 큐레이팅이라는 점을 명백히 하고 있다(Choudhury and Lynch, 2009). 이러한 변화는 도서관이 정보를 위한 보존 공간이 아니라 아이디어와 지식을 공유하기 위한 기회와 인프

라를 제공하는 주체 및 공간으로서 위상 변화를 요구하고 있다. 특히, 연구 및 대학도서관은 도서관을 직접 방문하는 이용자가 감소하는 경향이 두드러진 가운데 이용자들에게 접근하는 새로운 방법을 찾아야 하는 상황에서 어떤 서비스를 제공해야 하는 지에 초점을 맞추어야 한다. 그러나 2007년까지 연구도서관의 디지털 큐레이션에 대한 관심은 대부분 디지털화된 도서관의 장서와 전자저널의 보존에 집중되었고, 데이터 관리 및 보존에 대한 관심은 상대적으로 미미하여 도서관의 주요 관심사가 아니었다(Gold, 2007). 즉 2007년 이전에는 이론적 측면에서 디지털 큐레이션의 의미와 도서관과 사서의 역할 등에 대한 근본적인 논의가 주로 이루어졌으나, 이후 연구 및 대학도서관을 중심으로 하는 데이터 큐레이션에 관한 연구와 교육이 중점적으로 진행되면서 도서관이 디지털 큐레이션을 위한 장기적인 공간이 될 것을 제안하는 많은 논의가 있었다. 미국에서는 존스홉킨스, 코넬, 테네시, 캘리포니아, 뉴멕시코 등 주요 대학들의 도서관 시스템과 연구자들이 미국 과학재단(NSF)의 DataNet 사업의 일환으로 지원되는 데이터 큐레이션 프로젝트에서 연구개발 파트너로 주도적인 역할을 한 바 있다. 또한 이 기간 동안 여러 문헌정보학 대학원에서 디지털 데이터의 큐레이션을 지원하기 위한 기술과 지식의 개발을 위한 새로운 교육사업을 시작하였다(Gold, 2010). 영국 DCC의 여러 디지털 큐레이션 프로젝트에서도 도서관이 주도적으로 참여한 바 있으며, 도서관이 운영하는 기관 리포지토리와 국가 또는 국제적인 리포지토리 간의 연계를 통해 향후 10년 동안 디지털 데이터 큐레이션이 도서관의 새로운 정보서비스로서 급

속히 발전하게 될 것으로 전망하였다. 특히, 데이터 관리(stewardship)와 전체 생애주기 동안의 정보 관리는 연구 경쟁력에 필수적이 되고 있으며, 다양한 방식으로 미래의 연구도서관은 현재 연구 사업에 필수적인 데이터와 지식을 수집, 보존, 전달하게 될 것이라고 전망하였다(McCue and Corson-Rikert, 2007).

이와 같이 정보관리 및 서비스에 대한 도서관의 전통적인 사명은 디지털 데이터로 확장되고 있으며, 디지털 정보의 보존을 넘어 시간이 흐른 후에 이들의 선정, 공유, 재사용이 가능함은 물론 데이터 마이닝, 데이터-시각화 등의 방법들을 통해 잘 큐레이팅된 데이터는 그 자체가 과학의 프로세스와 실행에서 새롭고 필수적이며 유용한 것이 되고 있다(Gold, 2010). 이러한 환경에서 특히 연구 및 대학도서관이 연구데이터 큐레이션 분야에 기여할 수 있는 새로운 영역들이 다음과 같이 제안되었다. 첫째, 스몰 사이언스 데이터 큐레이팅(Small Science Data Curating). 국제적·국가적인 사업을 통해 기상, 천문, 유전학 등의 빅 사이언스(Big Science) 연구에서 생성되는 대규모의 데이터(빅 데이터/Big Data)와는 달리 개인이나 소규모 집단이 수행하는 스몰 사이언스 연구에서 생산되어 개인 컴퓨터에서 사용 및 조작이 가능한 소규모 또는 중규모(스몰데이터/Small Data) 데이터에 대한 큐레이팅을 제공하여야 한다(Carlson, 2006; Gold, 2007; Purdue University Libraries, 2007; Choudhury, 2008). 둘째, 데이터 출판 지원. 연구성과물로서 출판되는 학술논문이나 단행본에 관련된 데이터의 출판 지원(Gold, 2007; Joint Task Force on Library Support for E-Science, 2007; Lynch, 2006). 즉, 연구성과물이 생산되는 과정에서 수

집·가공·생산된 데이터의 출판을 지원할 수 있어야 하며, 이를 위해서는 데이터 관련 메타데이터와 데이터 인용 등에 대한 지식과 역량이 필요하다. 셋째, 주제 리포지토리. 특정 분야 주제 리포지토리들의 호스트로서 주도적인 역할을 담당할 수 있어야 한다(Messerschmitt, 2003; Steinhart, Saylor, & Westbrooks, 2008).

### 2.3 디지털 큐레이션 연구 영역

디지털 큐레이션 연구 영역을 파악하기 위해 해당 주제를 다룬 다양한 문헌들을 검토하였다. <표 1>은 관련 문헌들을 검토한 결과로 디지털 큐레이션 연구 영역을 크게 세 가지로 구분한 것이다. <표 1>에서 보는 바와 같이 디지털 큐레이션 연구 영역은 크게 이론 및 인프라, 디지털 정보관리 및 서비스, 기관의 3개 범주로 구분할 수 있다. 첫째, 이론 및 인프라 영역은 디지털 정보관리 및 서비스의 토대가 되는 이론 및 모델, 정책 및 전략, 국제·국가적 사업, 포맷 및 자료유형 등의 주제가 포함된다. 둘째, 정보관리 및 서비스 영역은 디지털 큐레이션의 실행에 해당하는 것으로 선정 및 수집, 접근 및 재사용, 저장과 보존 및 아카이빙, 메타데이터, 저작권을 포함하는 권한 관리 및 보안, 표준(파일 포맷, 인코딩 등), 훈련·교육, 지침·매뉴얼, 홍보, 모범실무, 연구(계획, 출판, 인용 등), 부가서비스(기획 및 분석 등)가 포함된다. 셋째, 기관 영역은 디지털 큐레이션의 현장으로서 도서관 및 리포지토리, 기록관, 박물관 관련 주제들이다.



〈표 1〉 문헌정보학 분야 디지털 큐레이션 연구 영역

| Lee & Tibbo (2007)      | Bailey (2012)                           | Jones et al. (2013)           | Higgins (2008) | Fearon et al. (2013) | 중분류                          | 대분류        |
|-------------------------|---|-------------------------------|----------------|----------------------|------------------------------|------------|
| -                       | 모델 및 정책, 전략, 국가·국제적 사업 및 기관 사례, 포맷 및 자료 | 연구데이터관리 정책 및 전략, 사업계획 및 지속가능성 | 개념화            | -                    | 이론(모델, 정보유형 등), 정책 및 전략, 인프라 | 이론 및 인프라   |
| -                       | -                                       | -                             | -              | -                    | 관리 일반 (유형별 관리)               | 정보관리 및 서비스 |
| 선정 및 평가                 | -                                       | 현용 데이터 관리, 데이터 선정 및 수집        | 평가, 선정, 폐기     | -                    | 선정 및 수집                      |            |
| 접근 및 재사용                | -                                       | -                             | 접근 및 재사용       | 데이터 공유와 접근           | 접근 및 재사용                     |            |
| 저장, 장기보존                | -                                       | -                             | 저장, 이관 및 보존    | 데이터 저장 및 백업          | 저장, 보존 및 아카이빙                |            |
| -                       | 저작권                                     | -                             | -              | 데이터 권한 관리, 보안 및 익명화  | 권한 관리 및 보안                   |            |
| 메타데이터                   | 메타데이터                                   | 데이터 목록                        | 메타데이터          | 메타데이터 선정 및 적용        | 메타데이터                        |            |
| 개방형 표준 (파일 포맷, 데이터 인코딩) | -                                       | -                             | -              | 파일포맷 결정, 파일명 부여와 정리  | 파일 포맷, 인코딩 등                 |            |
| 정보관리 리터러시 촉진            | -                                       | 지침, 훈련 및 지원                   | -              | 모범실무 교육              | 교육 및 훈련, 지침/매뉴얼, 홍보          |            |
| -                       | 연구                                      | 데이터 관리 계획                     | -              | 데이터관리계획작성 지원 데이터 인용  | 연구(계획, 출판, 인용 등)             |            |
| -                       | 서비스                                     | -                             | -              | -                    | 부가서비스 (기획, 분석 등)             |            |
| 디지털 리포지토리               | 디지털 리포지토리                               | 데이터 리포지토리                     | -              | 데이터 리포지토리, 기관 리포지토리  | 도서관 리포지토리, 기록관, 박물관 등        | 기관         |

### 3. 디지털 큐레이션 연구동향

#### 3.1 데이터 및 방법

디지털 큐레이션에 관한 연구동향을 문헌정보학 분야를 중심으로 살펴보았다. 이를 위해 문헌정보학 분야의 대표적인 학술 데이터베이스의 하나인 LISTA에서 디지털 또는 데이터 큐레이션을 다룬 학술논문을 검색하여 문헌집단을 구성하였다. 2014년 7월에 LISTA에서 검색한 관련 논문은 총 376편이었으나, 디스크립

터 프로파일링(김관준, 이재윤, 2007)을 위한 필수적인 레코드로서 디스크립터 또는 키워드(제목과 초록) 필드의 값이 없는 2개 논문을 제외한 총 374편의 학술논문을 지적구조 분석에 사용하였다. 실제 검색을 위한 탐색전략은 레코드 내 텍스트에서 '(digital or data) and curat\*'이 출현하고, 검색시점까지 전체 기간 동안 영어로 작성된 학술논문으로 제한하였다. 그 결과 분석에 사용된 문헌집합의 디스크립터와 키워드 통계는 〈표 2〉와 같다.

이러한 374개 논문의 논문번호, 출판년도, 디

〈표 2〉 문헌집합 통계

| 항 목                            | 내 역         |
|--------------------------------|-------------|
| 전체 논문 수                        | 374         |
| 디스크립터 문헌빈도(최소/최대/평균)           | 1/71/2.64   |
| 키워드 문헌빈도(최소/최대/평균)             | 1/283/6.10  |
| 논문 당 디스크립터 수(최소/최대/평균)         | 1/23/6.53   |
| 논문 당 키워드 수(최소/최대/평균)           | 11/284/88.8 |
| 전체 디스크립터 종수/분석 디스크립터 종수(df)=7) | 926/63      |
| 전체 키워드 종수                      | 3666        |

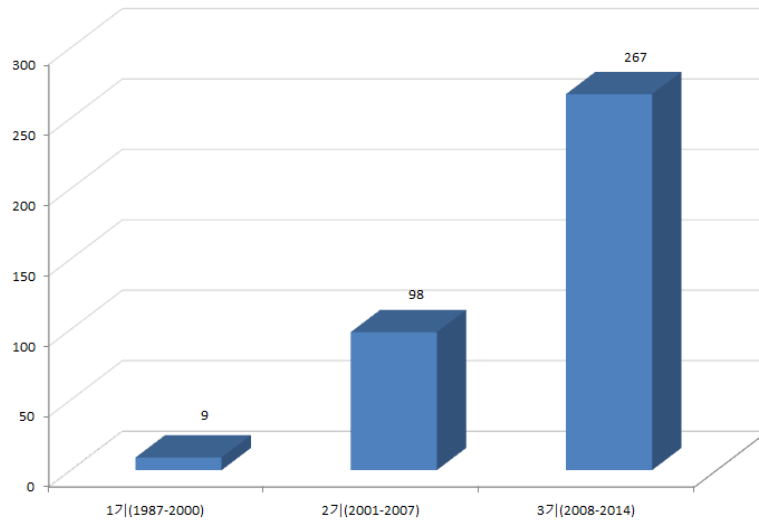
스크립터, 제목, 초록 필드를 대상으로 연구 동향을 분석하였다. 특히, 디스크립터 프로파일링은 각 논문에 부여된 평균 디스크립터 수에 기초하여 분석대상 기간(1987년~2014년)인 28년간 7회 이상 출현한 것으로 문헌빈도가 7 이상인 상위 63개 디스크립터를 대상으로 하였다. 주요 분석내용은 첫째, 기본 분석으로 문헌정보학 분야 디지털 큐레이팅 영역의 시기별 논문수, 시기별 초출 디스크립터의 추이를 살펴보고, 둘째, 지적구조 분석으로 디스크립터 프로파일링에 기초하여 시기별·영역별 연구지도 생성, 클러스터 분석, 네트워크 중심성 지수, 생산지수와 성장지수를 적용한 결과로 구체적인 연구 동향을 파악하였다. 여기서 문헌집합에 대한 사전처리 및 디스크립터 프로파일링을 위한 프로그램은 Python 및 Visual FoxPro로 구현된 프로그램과 SPSS 통계 프로그램을 사용하였고, 네트워크 중심성 지수 산출 및 클러스터 분석, 연구지도 생성은 Wnet(이재운, 2013)과 NodeXL을 사용하였다.

### 3.2 기본 분석

LISTA 데이터베이스에서 디지털 큐레이션

관련 논문이 처음 출현한 1987년부터 검색시점인 2014년 7월까지 28년을 각각 14년, 7년, 7년의 3개 시기로 구분하여 살펴본 시기별 논문 수는 〈그림 1〉과 같다. 이와 같이 3개 시기로 구분한 이유는 이전의 14년 동안(1987년~2000년)은 관련 논문이 9개로 미미한 수준이며, 2001년 이후 크게 논문 수가 증가하면서 2007년 전후로 상당한 변화가 발생하였기 때문이다(Gold, 2007; 2010).

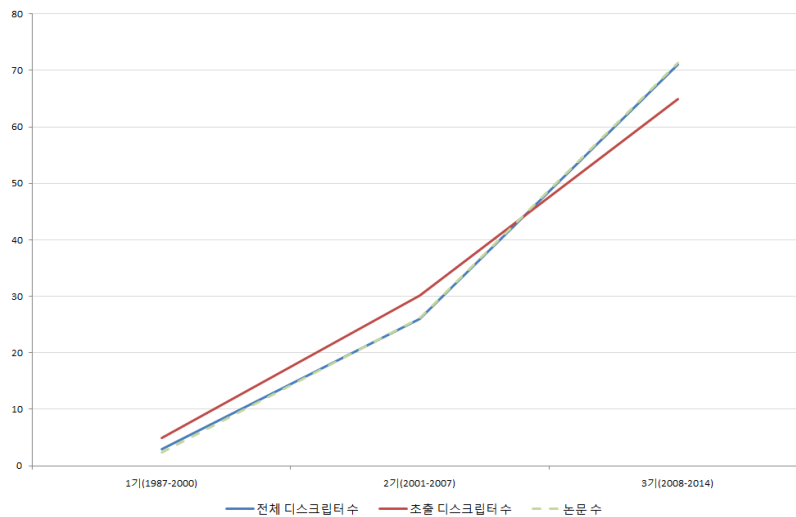
〈그림 1〉에서 보는 바와 같이 1기에 비해 2기(2001년~2007년)의 논문 수가 10배 이상 증가하였고, 이후 3기(2008년~2014년)에는 30배로 증가하고 있어 2000년대 이후 문헌정보학 분야의 디지털 큐레이션 관련 연구가 지속적으로 크게 성장하고 있음을 알 수 있다. 각 시기별 전체 디스크립터 수, 초출 디스크립터 수와 논문 수를 비교한 〈표 3〉과 이를 비율로 표시한 〈그림 2〉에서도 이러한 경향은 동일하게 나타났다. 즉, 2000년대 이후 문헌정보학 분야에서 디지털 큐레이션은 지속적으로 그 영역이 확장되고 있으며 새로운 연구 주제도 큰 폭으로 증가하고 있다는 것을 알 수 있다.



〈그림 1〉 문헌정보학 분야 디지털 큐레이션 영역의 시기별 논문 수

〈표 3〉 3개 시기별 전체 디스크립터 수, 초출 디스크립터 수, 논문 수

| 구분            | 전체 디스크립터 수 | 초출 디스크립터 수 | 논문 수 |
|---------------|------------|------------|------|
| 1기(1987-2000) | 49         | 46         | 9    |
| 2기(2001-2007) | 442        | 279        | 98   |
| 3기(2008-2014) | 1209       | 601        | 267  |
| 합계            | 1700       | 926        | 374  |



〈그림 2〉 3개 시기별 전체 디스크립터 수, 초출 디스크립터 수, 논문 수의 비율

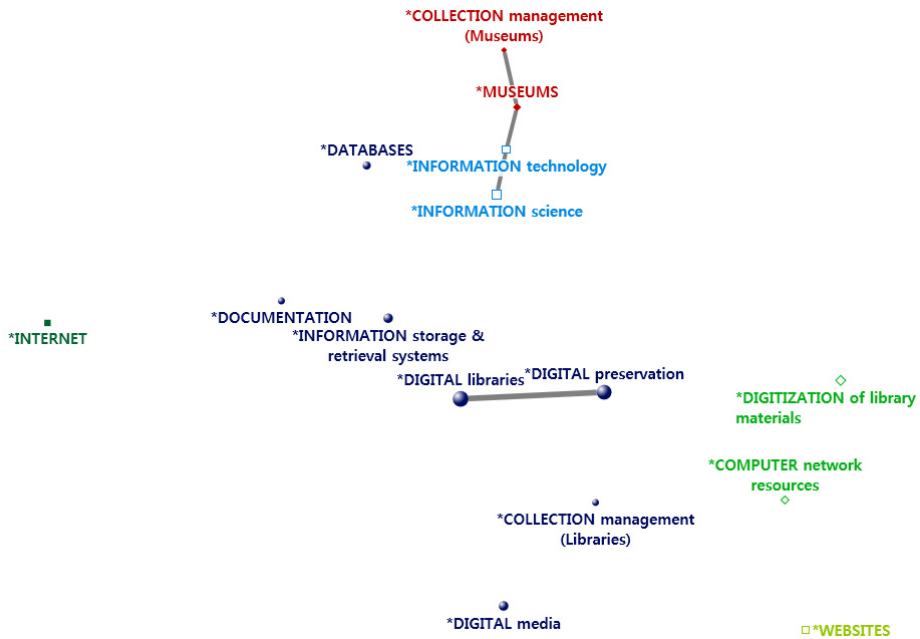
### 3.3 지적구조 분석

#### 3.3.1 시기별 분석

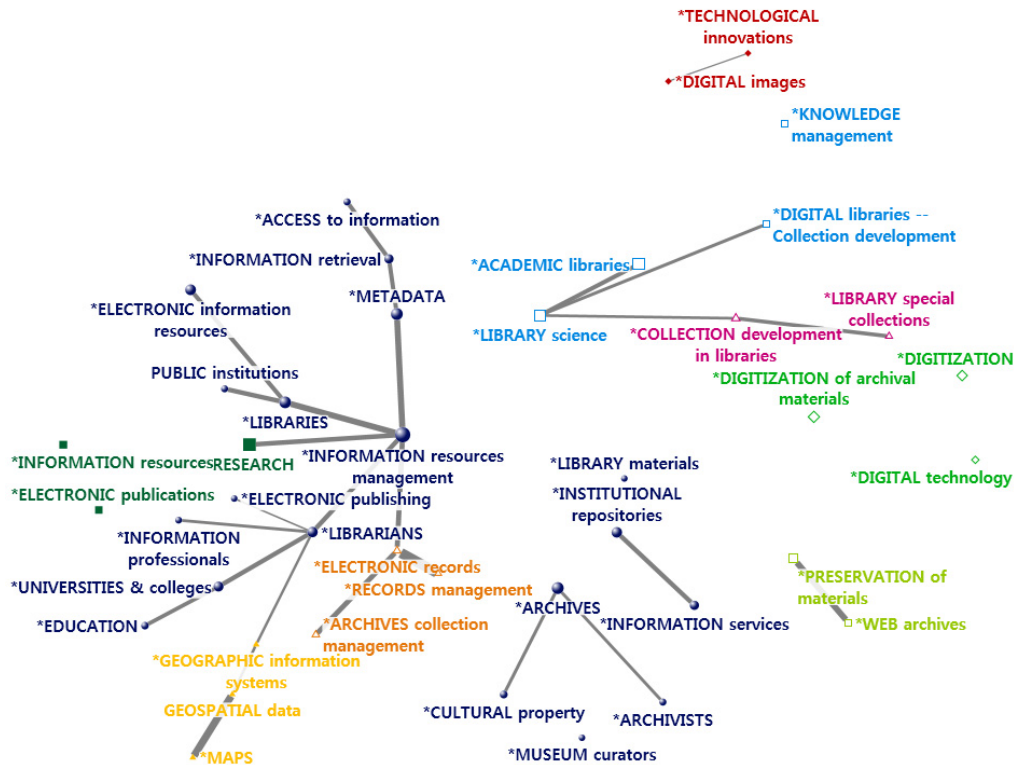
시기별 초출 디스크립터를 중심으로 연구지도를 생성하여 시기별 연구동향을 살펴보았다. 1기의 초출 디스크립터(16개)에 기초한 연구지도는 <그림 3>이며, 문헌정보학 분야에서 초기의 디지털 큐레이션 연구는 박물관과 디지털 도서관을 중심으로 시작되었다. 중앙 위쪽에 보이는 것처럼 박물관의 장서관리에 정보학과 정보 기술을 적용하는 연구가 있었다. 중앙 아래쪽에는 디지털 도서관을 중심으로 디지털 매체의 장서관리와 디지털 보존이, 좌측에는 정보검색, 다큐멘테이션, 데이터 분석, 데이터베이스에 관한 주제들이 출현하였다. 또한, 우측에는 도서관 자료의 디지털화, 컴퓨터 네트워크 자원, 웹

사이트 관련 주제가 출현하였다. 따라서 1기는 문헌정보학 분야에서 디지털 큐레이션에 관한 뚜렷한 개념이 없이 박물관과 도서관을 중심으로 기존의 정보관리 및 서비스에 정보기술을 부분적으로 적용하는 차원에서 연구가 이루어진 시기이다.

2기의 초출 디스크립터(43개)에 기초한 연구지도는 <그림 4>로서 중앙 좌측으로 정보자원 관리와 연구(\*RESEARCH)가, 중앙 위쪽으로는 도서관학을 매개로 대학 및 연구도서관 영역이 출현 및 연결되었다. 아래쪽으로는 기관 리포지토리와 기록관, 전자기록과 지리정보 등이 연결되어 디지털 큐레이션 관련 주제와 정보 유형의 확장이 진행되었다. 또한 정보자원관리가 처음 출현하여 메타데이터, 정보검색, 정보 접근 등 접근 측면의 연구가 시작되었고, 사서를



<그림 3> 시기별 연구지도: 1기(1987년~2000년)



〈그림 4〉 시기별 연구지도: 2기(2001년~2007년)

대상으로 하는 정보전문가 교육에 관한 연구가 출현하였다. 이외에도 장서관리, 디지털 보존 및 디지털화 등에 관한 다양한 연구가 수행되었다. 또한 이 시기에 디지털 도서관 측면에서 정보서비스가 출현하였고, 보존에서 아카이빙(\*WEB archives)로의 연계 및 확장이 있었다. 따라서 2기(2001년~2007년)의 디지털 큐레이션 연구는 정보관리 측면에서 관련 주체(기관, 전문직) 및 대상의 확장, 대학 및 연구도서관으로 세분화, 사서 중심의 전문직 교육, 정보서비스의 출현, 보존에서 아카이빙으로의 확장 등으로 요약할 수 있다. 즉, 2기는 정보관리 차원의 다양한 주체들이 본격적으로 출현 및 성장하는

가운데 정보서비스로의 전환이 일부 시작된 시기이다.

3기(2008년-2014년)의 초출 디스크립터(4개)에 기초한 연구지도는 〈그림 5〉이다. 이들 4개 디스크립터는 관리 측면의 2개(데이터베이스 관리, 관리)와 서비스 측면의 2개(교육과정, 큐레이터직(CURATORSHIP))으로 구분할 수 있다. 즉 최근의 3기(2008년~2014년)에는 이전 2기(2001년~2007년)의 연구 주체들이 대부분 그대로 유지되면서 소수의 새로운 주체(데이터베이스 관리, 관리, 교육과정, 큐레이터직)가 추가되었다. 여기서 유의할 것은 전체 시기 중에서 3기에 초출 디스크립터(전체 초출 디



〈그림 5〉 시기별 연구지도: 2기(2008년~2014년)

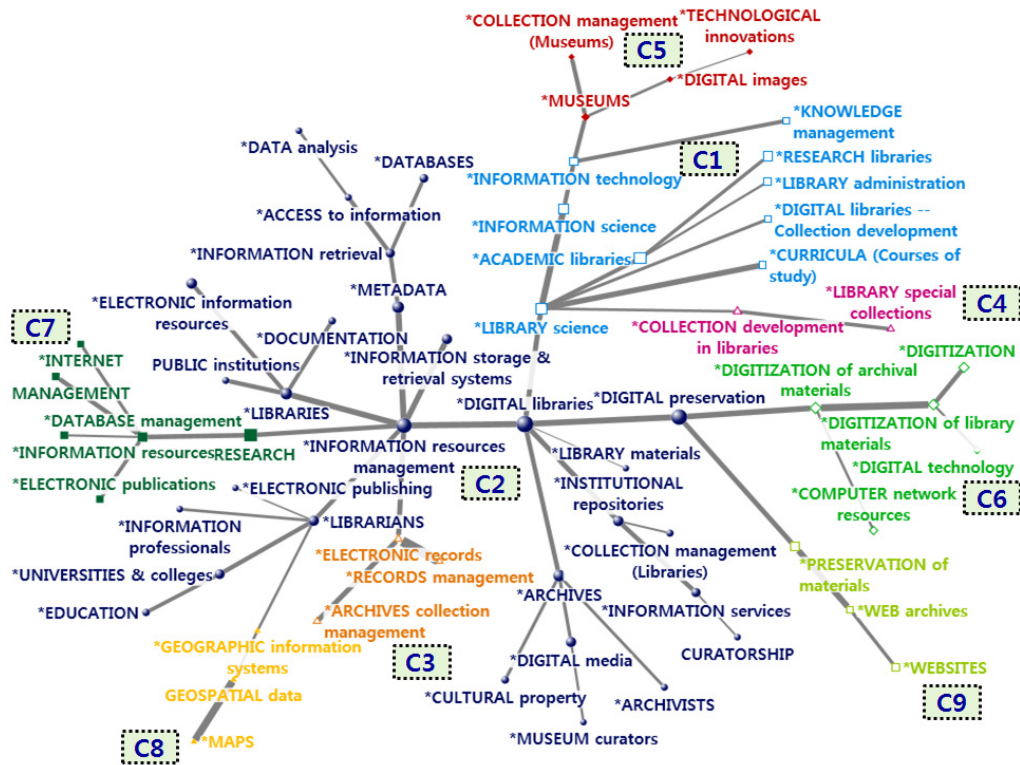
스크립터 926개 중 601개(64.9%))는 가장 많이 증가하였지만, 7년 동안 문헌빈도 7 이상의 기준을 충족한 것이 이들 4개라는 것이다(601개 중 4개(0.67%)).

### 3.3.2 영역별 분석

디지털 큐레이션 관련 학술논문 374편에 부여된 문헌빈도 7 이상의 63개 디스크립터는 문헌정보학 분야에서 연구되고 있는 주요 주제들이라고 할 수 있다. 〈그림 6〉은 NodeXL 프로그램을 이용하여 생성한 연구지도로서 문헌정보학 분야의 디지털 큐레이션 연구 동향을 주요 영역별로 나타낸 것이다. 여기서 각 노드는 63개 디스크립터(연구 주제)를 나타내고 노드의 크기는 해당 주제를 다룬 논문 수(문헌빈도)이다. 각 노드간의 링크(연결선)는 상호 연결된 노드들 간의 주제적 연관성이 크다는 것을 의미하고 링크의 굵기(폭)는 연관성의 정도를 나타낸다.

또한 동일한 모양으로 표시된 노드들은 클러스터 분석 결과에 따라 형성된 9개 군집을 표시한 것으로, 특정 군집에 속한 노드들은 함께 연구되는 경향이 많은 주제들로서 상호 관련성이 높은 주제들이라 할 수 있다.

〈그림 6〉에서 가장 두드러지는 영역은 중앙부의 디지털 도서관, 정보자원관리, 디지털 보존이 포함된 가장 큰 군집이다. 이러한 군집(C2)은 디지털 문화유산 기관으로서 디지털 도서관, 기관 리포지토리, 기록관의 정보서비스와 디지털 정보관리의 핵심 영역으로서 정보 접근 및 메타데이터에 관한 주제들과 사서를 중심으로 하는 전문가 교육 및 서비스 관련 주제들을 포함하고 있다. 중앙의 디지털 도서관의 상부에 있는 3개의 군집은 대학 및 연구도서관을 중심으로 하는 영역(C1)과 박물관을 중심으로 하는 영역(C5), 도서관의 특수장서 개발(C4)에 관한 영역이라 할 수 있다. 또한 우측의 디지털



〈그림 6〉 영역별 연구지도: 전체 시기(2008년~2014년)

보존에서 갈라지는 2개의 군집은 디지털화 관련 주제(C6)와 보존 및 웹아카이빙 관련 주제(C9)로 구성되어 있다. 좌측의 정보자원관리 아래쪽의 2개 클러스터는 기록관리(C3)와 지리정보 관련 영역(C8)이며, 맨 좌측의 클러스터는 연구 관련 영역(C7)에 해당한다. 이에 따라 문헌정보학 분야에서 디지털 큐레이션 연구는 디지털 문화유산 기관(도서관 및 기관 리포지토리, 기록관, 박물관)의 정보관리 및 보존을 중심으로 하고 있으며, 특히 해당 기관들의 장서관리와 보존에 관한 연구가 주류를 이루고 있음을 알 수 있다.

〈표 4〉는 〈그림 3〉의 63개 디스크립터들이 속

한 군집과 네트워크 중심성 지수 산출 결과(이재윤, 2006)이다. 여기서 전역 중심성(AVGSI)이 가장 큰 디스크립터는 정보자원관리(0.37), 디지털 도서관(0.37), 디지털 보존(0.35) 등이며, 이들 주제가 문헌정보학 분야의 디지털 큐레이션 연구의 주류가 되고 있음을 재차 확인할 수 있다. 〈표 4〉에서 각 클러스터를 대표하는 클러스터명은 해당 클러스터 내에서 문헌빈도와 3개 네트워크 중심성 지수(전역 중심성(AVGSI)과 지역 중심성(rNNC), 삼각매개 중심성(rTBC))를 더한 값(지수합)이 가장 큰 디스크립터들을 검토하여 부여하였다. 또한 〈그림 3〉에서 중앙의 수평부와 클러스터 간의 경계

〈표 4〉 문헌정보학 분야 디지털 큐레이션 연구 주제별 네트워크 중심성 지수

| 군집                     | 디스크립터  | 문헌빈도 | rTBC | AVGSIM | rNNC | 지수합         | 대표 주제명        |
|------------------------|--|------|------|--------|------|-------------|---------------|
| C1                     | *LIBRARY science                             | 26   | 0.70 | 0.34   | 0.06 | <b>1.10</b> | 대학 및<br>연구도서관 |
|                        | *ACADEMIC libraries                          | 30   | 0.64 | 0.32   | 0.03 | <b>0.99</b> |               |
|                        | *INFORMATION science                         | 21   | 0.56 | 0.31   | 0.03 | <b>0.90</b> |               |
|                        | *INFORMATION technology                      | 13   | 0.39 | 0.27   | 0.02 | 0.68        |               |
|                        | *RESEARCH libraries                          | 17   | 0.26 | 0.25   | 0    | 0.51        |               |
|                        | *KNOWLEDGE management                        | 8    | 0.23 | 0.26   | 0    | 0.49        |               |
|                        | *DIGITAL libraries -- Collection development | 7    | 0.16 | 0.23   | 0    | 0.39        |               |
|                        | *LIBRARY administration                      | 10   | 0.12 | 0.21   | 0    | 0.33        |               |
|                        | *CURRICULA (Courses of study)                | 7    | 0.09 | 0.22   | 0    | 0.31        |               |
| C2                     | *INFORMATION resources management            | 57   | 0.82 | 0.37   | 0.08 | <b>1.27</b> | 정보관리          |
|                        | *DIGITAL libraries                           | 71   | 0.82 | 0.37   | 0.08 | <b>1.26</b> |               |
|                        | *DIGITAL preservation                        | 62   | 0.75 | 0.35   | 0    | <b>1.09</b> |               |
|                        | *LIBRARIES                                   | 32   | 0.71 | 0.34   | 0.05 | 1.09        |               |
|                        | *ARCHIVES                                    | 33   | 0.71 | 0.33   | 0.05 | 1.09        |               |
|                        | *LIBRARIANS                                  | 25   | 0.67 | 0.33   | 0.05 | 1.04        |               |
|                        | *METADATA                                    | 33   | 0.62 | 0.32   | 0.02 | 0.96        |               |
|                        | *INSTITUTIONAL repositories                  | 25   | 0.57 | 0.32   | 0.03 | 0.92        |               |
|                        | *INFORMATION storage & retrieval systems     | 23   | 0.59 | 0.32   | 0    | 0.90        |               |
|                        | *UNIVERSITIES & colleges                     | 22   | 0.51 | 0.30   | 0.02 | 0.82        |               |
|                        | *DIGITAL media                               | 24   | 0.51 | 0.29   | 0.02 | 0.82        |               |
|                        | *ELECTRONIC information resources            | 26   | 0.50 | 0.30   | 0    | 0.80        |               |
|                        | *INFORMATION services                        | 18   | 0.46 | 0.30   | 0.02 | 0.78        |               |
|                        | *INFORMATION retrieval                       | 18   | 0.39 | 0.28   | 0.03 | 0.70        |               |
|                        | *EDUCATION                                   | 11   | 0.29 | 0.26   | 0    | 0.54        |               |
|                        | *ACCESS to information                       | 9    | 0.23 | 0.25   | 0.02 | 0.50        |               |
|                        | *INFORMATION professionals                   | 8    | 0.25 | 0.25   | 0    | 0.49        |               |
|                        | *CULTURAL property                           | 12   | 0.24 | 0.24   | 0    | 0.48        |               |
|                        | *DOCUMENTATION                               | 11   | 0.20 | 0.24   | 0    | 0.44        |               |
|                        | PUBLIC institutions                          | 9    | 0.20 | 0.23   | 0    | 0.44        |               |
|                        | *COLLECTION management (Libraries)           | 10   | 0.20 | 0.23   | 0    | 0.43        |               |
|                        | CURATORSHIP                                  | 8    | 0.19 | 0.23   | 0    | 0.41        |               |
|                        | *DATABASES                                   | 16   | 0.17 | 0.23   | 0    | 0.40        |               |
|                        | *ARCHIVISTS                                  | 9    | 0.13 | 0.22   | 0    | 0.35        |               |
| *DATA analysis         | 8  | 0.13 | 0.22 | 0      | 0.35 |             |               |
| *MUSEUM curators       | 7  | 0.07 | 0.20 | 0      | 0.27 |             |               |
| *ELECTRONIC publishing | 7  | 0.00 | 0.16 | 0      | 0.16 |             |               |
| *LIBRARY materials     | 7  | 0.01 | 0.15 | 0      | 0.16 |             |               |



| 군집 | 디스크립터                                | 문헌빈도 | rTBC | AVGSIM | rNNC | 지수합         | 대표 주제명 |
|----|--------------------------------------|------|------|--------|------|-------------|--------|
| C3 | *ELECTRONIC records                  | 12   | 0.34 | 0.28   | 0.03 | <b>0.65</b> | 기록관리   |
|    | *ARCHIVES collection management      | 10   | 0.21 | 0.25   | 0    | <b>0.47</b> |        |
|    | *RECORDS management                  | 7    | 0.09 | 0.24   | 0.02 | <b>0.34</b> |        |
| C4 | *COLLECTION development in libraries | 14   | 0.25 | 0.24   | 0.02 | <b>0.51</b> | 장서개발   |
|    | *LIBRARY special collections         | 9    | 0.06 | 0.19   | 0.02 | <b>0.26</b> |        |
| C5 | *MUSEUMS                             | 16   | 0.32 | 0.27   | 0.03 | <b>0.62</b> | 박물관    |
|    | *DIGITAL images                      | 8    | 0.15 | 0.22   | 0.02 | <b>0.39</b> |        |
|    | *TECHNOLOGICAL innovations           | 8    | 0.11 | 0.20   | 0    | <b>0.31</b> |        |
|    | *COLLECTION management (Museums)     | 7    | 0.03 | 0.18   | 0.02 | 0.22        |        |
| C6 | *DIGITIZATION of archival materials  | 29   | 0.59 | 0.32   | 0.03 | <b>0.94</b> | 디지털화   |
|    | *DIGITIZATION of library materials   | 25   | 0.56 | 0.31   | 0.05 | <b>0.92</b> |        |
|    | *DIGITIZATION                        | 24   | 0.37 | 0.27   | 0    | <b>0.64</b> |        |
|    | *COMPUTER network resources          | 14   | 0.19 | 0.23   | 0    | 0.42        |        |
|    | *DIGITAL technology                  | 10   | 0.17 | 0.23   | 0    | 0.40        |        |
| C7 | *RESEARCH                            | 37   | 0.69 | 0.33   | 0.02 | <b>1.04</b> | 연구     |
|    | *DATABASE management                 | 20   | 0.53 | 0.30   | 0.08 | <b>0.91</b> |        |
|    | MANAGEMENT                           | 11   | 0.38 | 0.28   | 0    | <b>0.66</b> |        |
|    | *ELECTRONIC publications             | 10   | 0.36 | 0.27   | 0    | 0.63        |        |
|    | *INTERNET                            | 8    | 0.29 | 0.25   | 0    | 0.55        |        |
|    | *INFORMATION resources               | 12   | 0.16 | 0.22   | 0    | 0.37        |        |
| C8 | *GEOGRAPHIC information systems      | 8    | 0.16 | 0.23   | 0    | <b>0.38</b> | 지리정보   |
|    | GEOSPATIAL data                      | 7    | 0.05 | 0.20   | 0.03 | <b>0.28</b> |        |
|    | *MAPS                                | 8    | 0.04 | 0.18   | 0.02 | <b>0.23</b> |        |
| C9 | *PRESERVATION of materials           | 13   | 0.30 | 0.28   | 0.02 | <b>0.59</b> | 웹아카이빙  |
|    | *WEB archives                        | 7    | 0.23 | 0.26   | 0.03 | <b>0.53</b> |        |
|    | *WEBSITES                            | 10   | 0.17 | 0.23   | 0    | <b>0.40</b> |        |

에 위치하면서 다른 군집이나 주제를 연결하고 있는 노드(디지털 도서관, 디지털 보존, 정보자원관리, 연구, 도서관학, 도서관, 아카이브, 메타 데이터)들은 삼각매개 중심성(rTBC) 값이 높은 마당밭에 해당하는 주제들이다.

### 3.3.3 종합 분석

문헌정보학 분야의 디지털 큐레이션 연구 동향을 생산성과 성장성 측면에서 종합적으로 살

펴보았다. 첫째, 생산성 측면에서 63개 주제들의 출판 논문 수에 기초하여 산출한 생산지수로 주요 영역들이 얼마나 많이 연구되었는지를 살펴 보았다. 둘째, 성장성 측면에서 디지털 큐레이션 관련 63개 주제들의 시기별 성장률에 기초하여 산출한 성장지수로 주요 영역들이 성장하고 있는 정도를 알아보았다.

생산지수는 특정 디스크립터가 부여된 논문 수( $DF_d$ )를 전체 디스크립터가 부여된 논문 수

(Total DF)로 나눈 백분율로서 산출 공식은 다음과 같다.

$$\text{생산지수} = \left( \frac{DF_{d_i}}{\text{Total DF}} \right) \times 100$$

성장지수는 1차 성장률(1기~2기)과 2차 성장률(2기~3기)을 산출하여 이들 2개 시기의 성장률을 평균한 시기별 평균 성장률로서 산출 공식은 다음과 같다. 여기서 이전 시기의 논문 수가 0인 경우는 성장률 계산이 불가능하므로, 이전 시기에도 해당 디스크립터의 개념은 존재한 것으로 전제하여 분모(이전시기 논문 수)에 1을 더한 보정 성장률을 사용하였다.

$$\text{성장지수} = \frac{(\text{1차 성장률} + \text{2차 성장률})}{2}$$

$$\text{성장률} = \left( \frac{(\text{해당시기 논문 수} - \text{이전시기 논문 수})}{\text{이전시기 논문 수}} \right) \times 100$$

$$\text{보정 성장률} = \left( \frac{(\text{해당시기 논문 수} - \text{이전시기 논문 수})}{\text{이전시기 논문 수} + 1} \right) \times 100$$

<표 5>는 문헌정보학 분야 디지털 큐레이션 관련 주요 연구 영역별 생산지수와 성장지수를 산출한 결과이다. 생산성 측면에서 문헌정보학 분야 디지털 큐레이션 관련 주요 연구 영역의 생산지수에 기초한 분석 결과는 다음과 같다. 문헌정보학 분야에서 디지털 큐레이션과 관련하여 가장 많이 연구된 주제는 디지털 도서관, 디지털 보존, 정보자원관리의 3개 영역이다. 즉, 문헌정보학 분야에서는 지금까지 디지털 도서관을 중심으로 기관 리포지토리와 기록관의 정보관리와 보존 측면의 연구가 가장 많이 이루어졌다. 특히, 대학 및 연구도서관을 중심으로

정보관리 및 서비스 측면에서 연구, 정보 접근 및 이용(메타데이터, 정보검색, 정보 접근), 디지털화 등이 활발하게 연구되고 있다.

성장성 측면에서 문헌정보학 분야 디지털 큐레이션 관련 주요 연구 영역의 성장지수에 기초한 분석 결과는 다음과 같다. 성장지수가 높게 나타나고 있는 것은 대부분 정보관리 및 기관 영역의 주제이다. 또한 서비스 영역에 속하는 일부 주제들(정보 접근, 정보 저장 및 검색, 연구, 전자출판 등)의 성장세도 높게 나타나고 있다. 여기서 주목할 점은 최근 3기(2008년~2014년)에 새로 출현한 데이터베이스 관리, 관리, 큐레이터직, 교육과정의 4개 주제가 모두 상당히 높은 성장을 보이고 있으며, 특히 큐레이터직과 교육과정은 서비스 영역에 해당한다. 따라서 문헌정보학 분야의 디지털 큐레이션 연구는 현재까지는 대부분 개별 기관들의 정보관리를 중심으로 하는 가운데 최근에 문화유산과 데이터 관리로의 확장이 활발하게 진행되고 있으며, 큐레이터직과 교육과정에 관한 연구가 크게 성장하고 있다. 이에 반해 디지털 큐레이션 관련 초기의 연구에 해당하는 박물관이나 장서관리, 지리정보 등의 주제는 상대적으로 성장지수가 낮은 것으로 나타났다. 여기서 생산성과 성장성 양자에서 디지털 큐레이션의 대상이 되는 정보자원의 개념과 실행 주체로서 기관의 확장 추세가 뚜렷함을 알 수 있다. 즉, 문헌정보학 분야의 디지털 큐레이션 연구는 지금까지 개별 기관의 유형별 정보관리가 주류를 이루고 있으나, 최근에는 디지털 큐레이션의 대상(문화유산 및 데이터)과 주체(문화유산기관)의 확장 및 통합 추세와 함께 그 중심이 점진적으로 정보관리에서 서비스로 전환되는 과정에 있다.

〈표 5〉 문헌정보학 분야 디지털 큐레이션 관련 연구 영역별 생산지수와 성장지수

| DID | 디스크립터                                    | 생산지수         | 1차 성장률 | 2차 성장률 | 성장지수        |
|-----|--|--------------|--------|--------|-------------|
| 128 | *DIGITAL libraries                       | <b>412.8</b> | 2100   | 118.2  | 1109.1      |
| 118 | *DATABASE management                     | 116.3        | 0      | 2000   | <b>1000</b> |
| 203 | *INFORMATION resources management        | <b>331.4</b> | 1900   | 100    | 1000        |
| 137 | *DIGITAL preservation                    | <b>360.5</b> | 1700   | 138.9  | 919.4       |
| 29  | *ARCHIVES                                | 191.9        | 1500   | 20     | 760         |
| 252 | *LIBRARIES                               | 186.0        | 1400   | 28.6   | 714.3       |
| 162 | *ELECTRONIC information resources        | 151.2        | 1300   | 0      | 650         |
| 330 | *METADATA                                | 191.9        | 1000   | 130    | 565         |
| 699 | MANAGEMENT                               | 64.0         | 0      | 1100   | <b>550</b>  |
| 101 | *CULTURAL property                       | <b>69.8</b>  | 100    | 1000   | <b>550</b>  |
| 219 | *INFORMATION storage & retrieval systems | 133.7        | 1100   | -16.7  | 541.7       |
| 386 | *RESEARCH                                | 215.1        | 500    | 540    | 520         |
| 141 | *DIGITIZATION of archival materials      | 168.6        | 800    | 162.5  | 481.3       |
| 140 | *DIGITIZATION                            | 139.5        | 800    | 100    | 450         |
| 4   | *ACADEMIC libraries                      | <b>174.4</b> | 500    | 400    | <b>450</b>  |
| 33  | *ARCHIVES collection management          | 58.1         | 100    | 800    | 450         |
| 272 | *LIBRARY administration                  | 58.1         | 100    | 800    | 450         |
| 297 | *LIBRARY science                         | 151.2        | 700    | 171.4  | 435.7       |
| 248 | *LIBRARIANS                              | 145.3        | 700    | 157.1  | 428.6       |
| 388 | *RESEARCH libraries                      | <b>98.8</b>  | 200    | 650    | <b>425</b>  |
| 229 | *INSTITUTIONAL repositories              | 145.3        | 400    | 425    | 412.5       |
| 569 | CURATORSHIP                              | 46.5         | 0      | 800    | 400         |
| 14  | *ACCESS to information                   | 52.3         | 100    | 700    | <b>400</b>  |
| 142 | *DIGITIZATION of library materials       | 145.3        | 700    | 100    | 400         |
| 422 | *UNIVERSITIES & colleges                 | 127.9        | 600    | 166.7  | 383.3       |
| 135 | *DIGITAL media                           | 139.5        | 600    | 128.6  | 364.3       |
| 102 | *CURRICULA (Courses of study)            | 40.7         | 0      | 700    | <b>350</b>  |
| 212 | *INFORMATION services                    | 104.7        | 600    | 100    | 350         |
| 206 | *INFORMATION retrieval                   | 104.7        | 500    | 160    | 330         |
| 233 | *INTERNET                                | 46.5         | -100   | 700    | 300         |
| 167 | *ELECTRONIC records                      | 69.8         | 200    | 400    | 300         |
| 207 | *INFORMATION science                     | 122.1        | 400    | 200    | 300         |
| 166 | *ELECTRONIC publishing                   | 40.7         | 100    | 500    | 300         |
| 285 | *LIBRARY materials                       | 40.7         | 100    | 500    | 300         |
| 70  | *COLLECTION development in libraries     | 81.4         | 400    | 150    | 275         |
| 341 | *MUSEUMS                                 | 93.0         | 500    | 50     | 275         |
| 202 | *INFORMATION resources                   | 69.8         | 500    | 40     | 270         |
| 367 | *PRESERVATION of materials               | 75.6         | 400    | 125    | 262.5       |
| 138 | *DIGITAL technology                      | 58.1         | 200    | 300    | 250         |
| 165 | *ELECTRONIC publications                 | 58.1         | 200    | 300    | 250         |
| 104 | *DATA analysis                           | 46.5         | 0      | 500    | 250         |

| DID | 디스크립터  | 생산지수 | 1차 성장률 | 2차 성장률 | 성장지수  |
|-----|--|------|--------|--------|-------|
| 797 | PUBLIC institutions                          | 52.3 | 500    | -20    | 240   |
| 153 | *EDUCATION                                   | 64.0 | 300    | 166.7  | 233.3 |
| 180 | *GEOGRAPHIC information systems              | 46.5 | 500    | -40    | 230   |
| 36  | *ARCHIVISTS                                  | 52.3 | 200    | 250    | 225   |
| 83  | *COMPUTER network resources                  | 81.4 | 200    | 233.3  | 216.7 |
| 304 | *LIBRARY special collections                 | 52.3 | 400    | 25     | 212.5 |
| 417 | *TECHNOLOGICAL innovations                   | 46.5 | 400    | 0      | 200   |
| 201 | *INFORMATION professionals                   | 46.5 | 200    | 200    | 200   |
| 437 | *WEBSITES                                    | 58.1 | 400    | -20    | 190   |
| 334 | *MUSEUM curators                             | 40.7 | 400    | -25    | 187.5 |
| 127 | *DIGITAL images                              | 46.5 | 300    | 66.7   | 183.3 |
| 243 | *KNOWLEDGE management                        | 46.5 | 300    | 66.7   | 183.3 |
| 318 | *MAPS  | 46.5 | 300    | 66.7   | 183.3 |
| 381 | *RECORDS management                          | 40.7 | 200    | 150    | 175   |
| 72  | *COLLECTION management (Libraries)           | 58.1 | 100    | 250    | 175   |
| 131 | *DIGITAL libraries -- Collection development | 40.7 | 300    | 33.3   | 166.7 |
| 429 | *WEB archives                                | 40.7 | 300    | 33.3   | 166.7 |
| 625 | GEOSPATIAL data                              | 40.7 | 300    | 33.3   | 166.7 |
| 120 | *DATABASES                                   | 93.0 | 250    | 0      | 125   |
| 225 | *INFORMATION technology                      | 75.6 | 50     | 166.7  | 108.3 |
| 73  | *COLLECTION management (Museums)             | 40.7 | 100    | 100    | 100   |
| 149 | *DOCUMENTATION                               | 64.0 | 100    | 25     | 62.5  |

#### 4. 디지털 큐레이션 연구과제

디지털 큐레이션은 현재 공학 분야는 물론 경영학, 인문학 등 다양한 학문 분야에서 연구가 활발히 진행되고 있으며, 문헌정보학 분야에서도 차세대 정보서비스로서 새롭게 부상하고 있는 영역이다. 디지털 정보의 융합 현상은 정보 관리 및 서비스의 대상과 주체는 물론 그 범위와 방법론에 이르기까지 근본적인 변화를 요구하고 있다. 특히, 문헌정보학은 전통적으로 정보 관리 및 서비스의 이론 및 실행을 위한 토대를 제공해 온 학문분야로서 이러한 변화에 능동적으로 대처하여 차세대 정보서비스로서 디지

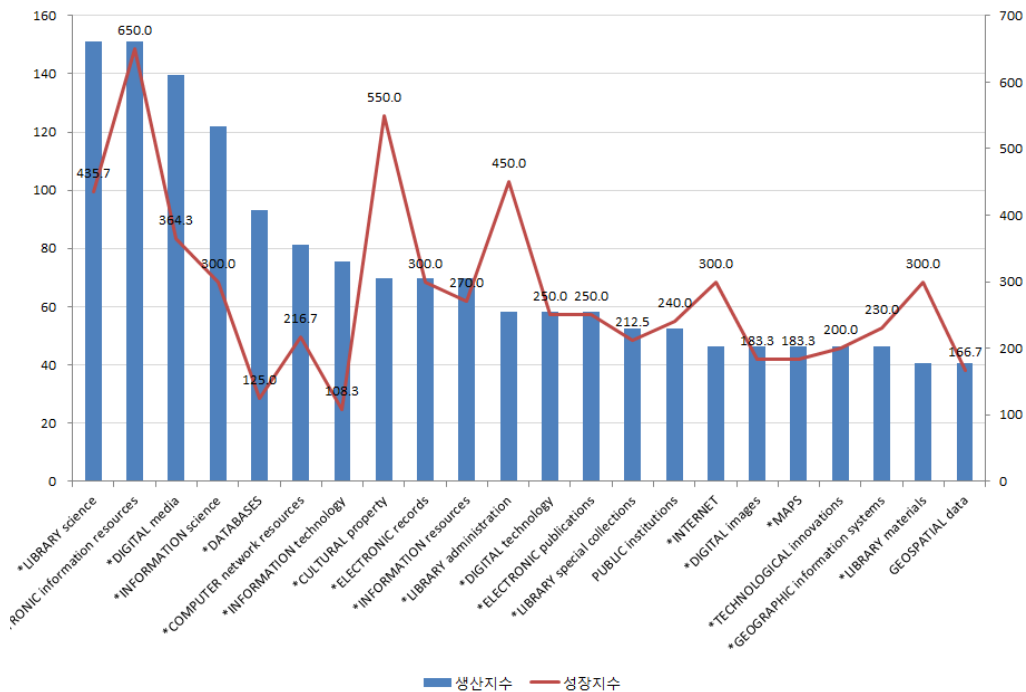
털 큐레이션을 주도할 책임을 갖는다. 따라서 지금까지 문헌정보학 분야의 디지털 큐레이션 연구동향을 분석한 결과에 기초하여, 향후 중점적으로 추진되어야 할 영역별 연구 과제를 제시하였다.

##### 4.1 이론 및 인프라

문헌정보학 분야 디지털 큐레이션 연구에서 이론 및 인프라 영역은 모델 및 정보 유형 등을 포함하는 이론, 정책 및 전략, 그리고 토대가 되는 정보기술 환경에 관한 주제들이다. <그림 7>은 이 영역에 해당하는 주제들의 생산지수와

성장지수를 표시한 것으로 생산성과 성장성 양자에서 이론 영역의 주제들이 전반적으로 높게 나타났다. 생산성 측면에서는 도서관학, 전자정보자원, 디지털 매체, 정보학, 데이터베이스, 컴퓨터 네트워크 자원 등의 주제들이, 성장성 측면에서는 전자정보자원, 문화유산, 도서관 경영 등이 높게 나타났다. 이를 통해 디지털 큐레이션의 이론적 토대를 제공하기 위한 도서관학, 정보학, 도서관 경영에 관한 연구와 정보서비스의 대상이 되는 정보 유형으로서 전자 또는 디지털 자원 등에 관한 연구가 활발히 이루어지고 있음을 알 수 있다. 특히, 성장성 측면에서는 이전의 도서관자료, 이미지, 지도 등 개별 정보 유형에 대한 연구보다 최근 출현한 문화유산에 관한 연구와 도서관 경영 관련 연구가 크게 성장

하고 있다. 그러나 아직은 서비스가 아닌 관리(도서관 경영)가 중심이 되고 있으며 서비스 모형 등 디지털 큐레이션의 이론적 토대를 제공할 수 있는 연구를 거의 찾아볼 수 없다. 또한, 인프라 영역의 정책 및 전략, 인프라(정보기술, 디지털 기술, 인터넷, 기술 혁신 등)에 관한 연구도 상대적으로 저조한 상황이다. 다시 말해서, 문헌정보학 분야 디지털 큐레이션 연구는 기존의 도서관학과 정보학 이론에 기초하여 디지털 자원에 정보기술을 적용하는 수준에 머물러 있으면서, 그 대상을 도서관자료, 전자기록, 데이터베이스 등을 아우르는 문화유산으로 확장하고 있음을 알 수 있다. 따라서 전통적인 아날로그 환경의 정보서비스를 단순히 확장하는 차원이 아닌 새로운 디지털 환경에서 차세대 디



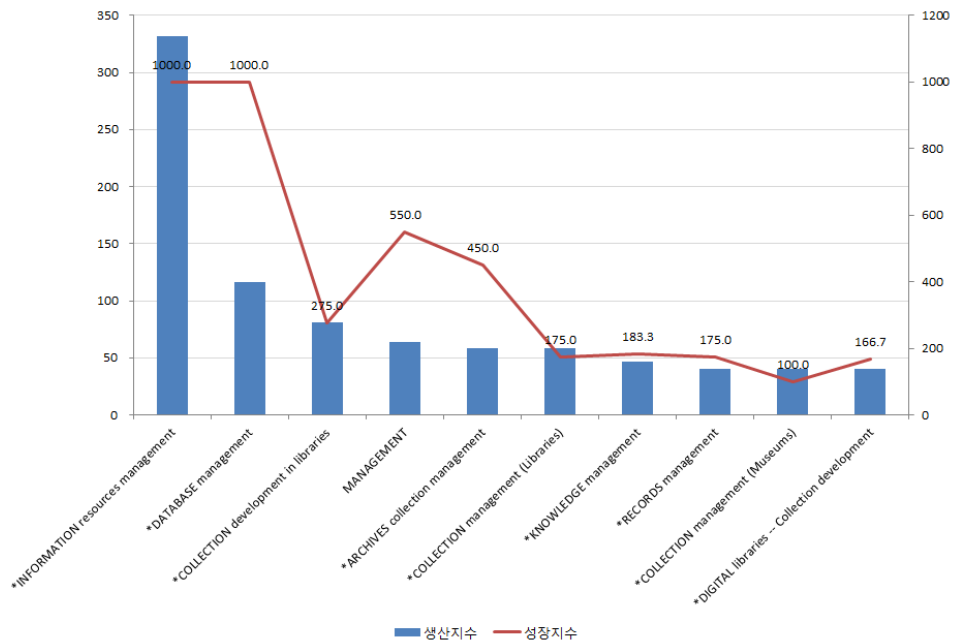
〈그림 7〉 이론 및 인프라 영역의 주제별 생산지수와 성장지수

지털 정보서비스로서 디지털 큐레이션을 위한 이론적 기반(이론, 모델 등)과 함께 이를 구체화하기 위한 정책 및 전략, 인프라(e-Research, e-Science, Cyberinfrastructure, platform 등)에 관한 연구는 현재까지 거의 이루어지지 않고 있다.

문헌정보학 분야의 디지털 큐레이션 연구에서 이론 및 인프라 영역의 연구과제로는 국가적 인프라의 구축과 연계하여 효율적인 디지털 큐레이션을 지원하기 위한 이론과 모형, 정책 및 전략에 관한 연구에 중점을 둘 필요가 있다. 특히, 이론 영역에서 각 기관의 개별적인 장서관리를 넘어 디지털 문화유산을 대상으로 새로운 정보관리 및 서비스를 추진하기 위한 디지털 큐레이션의 이론적 토대와 전략을 마련하는 연구는 시급하다 할 것이다.

### 4.2 정보관리 및 서비스

정보관리 및 서비스 영역은 디지털 큐레이션의 실질적인 내용에 해당하는 것으로 디지털 정보관리와 이를 기반으로 하는 서비스들이 포함된다. 이는 크게 세 가지로 구분하면 관리 일반(유형별 관리 포함), 관리 중심의 영역(선정 및 수집, 접근 및 재사용, 저장/보존 및 아카이빙, 권한 관리 및 보안, 메타데이터 등), 서비스 중심의 영역(교육 및 훈련, 지침/매뉴얼, 홍보, 연구(계획, 출판, 인용 등), 부가서비스(기획 및 분석 등)으로 세분할 수 있다. 먼저 관리 일반(유형별 관리 포함) 영역에서는 <그림 8>에서 보는 바와 같이 생산성 측면에서 정보자원 관리, 데이터베이스 관리, 도서관 장서관리, 관리의 순으로 높게 나타났다. 성장성 측면에서

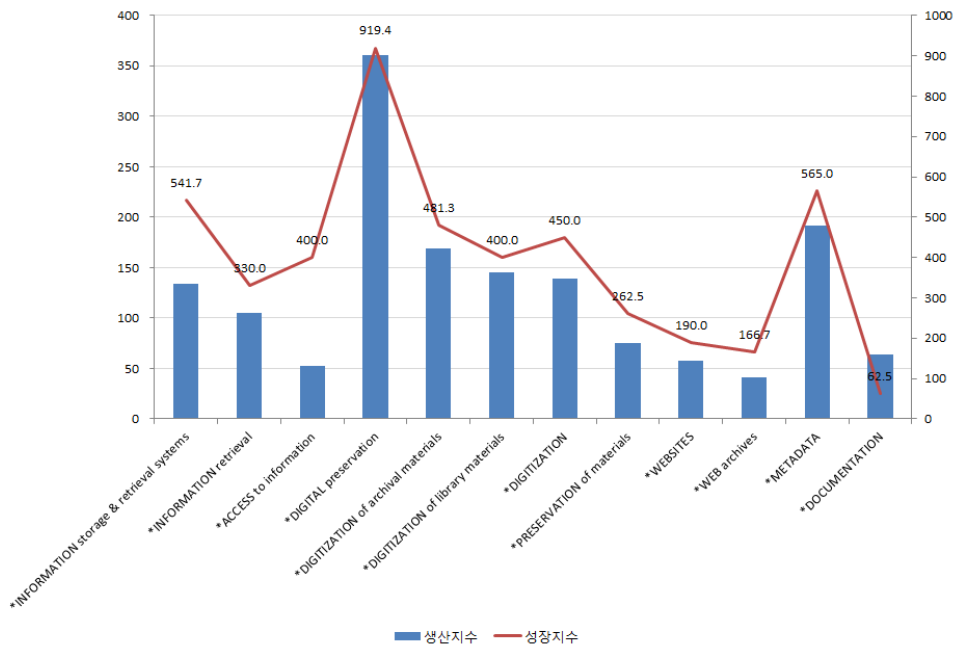


<그림 8> 관리 일반 영역의 주제별 생산지수와 성장지수

는 정보자원관리, 데이터베이스 관리, 기록관 장서관리, 관리의 순인 것으로 나타났으며, 특히 성장성 측면에서 최근 출현한 데이터베이스 관리의 성장이 두드러졌다. 그러므로 관리 일반 영역에서는 이전의 개별 기관의 장서관리(도서관, 기록관, 박물관)에서 보다 일반적인 정보자원관리 측면의 연구가 증가하였으며, 특히 데이터베이스 관리에 대한 연구가 최근에 크게 성장하고 있음을 알 수 있다. 관리 일반 영역의 향후 연구 과제로는 디지털 큐레이션의 대상이 되는 정보자원으로서 디지털 문화유산의 전반적인 관리에 관한 연구와 함께 새로운 정보자원인 데이터 관리에 관한 연구가 중점적으로 이루어져야 할 것이다.

〈그림 9〉는 정보관리 영역의 생산성과 성장성을 나타낸 것이다. 이 영역은 전반적으로 생

산성과 성장성이 높아 현재 디지털 큐레이션 분야 연구에서 정보관리 측면의 연구가 주류를 이루고 있음을 알 수 있다. 특히 디지털 보존과 메타데이터, 디지털화 관련 주제들은 생산성과 성장성이 가장 높아 해당 영역 연구를 주도하고 있다. 또한 정보저장 및 검색과 정보검색은 사실상 거의 동일한 의미이며 시간의 경과에 따라 전자에서 후자로 전환된 디스크립터라 할 수 있으므로, 정보검색(정보저장과 검색 포함)과 정보 접근은 생산성에 비해 성장성이 가장 큰 주제라고 할 수 있다. 반면, 다큐멘테이션과 메타데이터는 생산성에 비해 성장성이 크지 않아 상대적으로 아직 연구가 저조한 주제이다. 이외에 보존 및 디지털화 관련 주제들이 활발히 연구되고 있는 반면, 아카이빙 관련 연구는 다소 침체되어 있다. 따라서 문헌정보학 분야



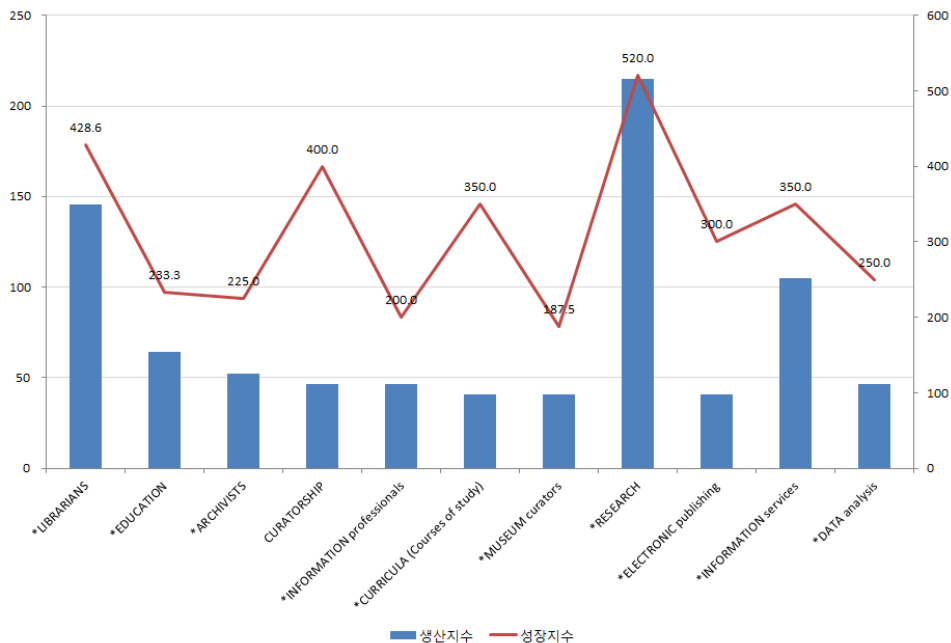
〈그림 9〉 정보관리 영역의 주제별 생산지수와 성장지수

에서 정보관리 영역의 연구는 보존, 디지털화, 정보 접근을 중심으로 성장하고 있으며, 메타데이터와 아카이빙 관련 연구는 다소 정체되어 있다.

정보관리 영역의 향후 연구과제는 현재 일부 주제(보존, 디지털화 등)에 치중된 경향을 벗어나 가치 있는 디지털 문화유산의 선정 및 수집, 재사용, 아카이빙, 권한 관리 및 보안(저작권 포함), 메타데이터, 파일 포맷 및 인코딩 등 다양하고 실질적인 주제들을 다루어야 한다. 특히, 디지털 큐레이션이 디지털 문화유산에 대한 새로운 정보서비스라는 인식에 기초하여 이용자에게 보다 효율적인 큐레이팅 서비스를 제공하기 위한 필수적인 단계로서의 정보관리를 전제로 연구가 이루어 져야 할 것이다. 또한, 현재와 같이 관리 지향의 보존 및 수집 중심이

아닌 서비스 측면의 아카이빙과 큐레이팅을 위한 연구가 필요하며, 이와 같은 맥락에서 그 토대가 되는 메타데이터에 관한 연구도 활성화되어야 한다.

〈그림 10〉은 문헌정보학 분야의 디지털 큐레이션 관련 연구에서 궁극적으로 지향하여야 하는 정보서비스 중심의 주제들에 대한 생산지수와 성장지수이다. 이에 해당하는 연구(계획, 출판, 인용 등), 부가서비스(기획 및 분석 등), 교육 및 훈련 등은 전반적으로 정보관리 중심의 영역과 비교하여 생산성이 낮은 반면 성장성은 높은 것으로 나타났다. 즉, 이들은 아직까지는 많이 연구가 이루어지고 있지 않으나 향후 발전 가능성이 높은 주제들이라 할 수 있으며, 이를 통해 디지털 큐레이션 연구가 점진적으로 정보관리 및 서비스(관리 중심)에서 정보



〈그림 10〉 정보서비스 영역의 주제별 생산지수와 성장지수

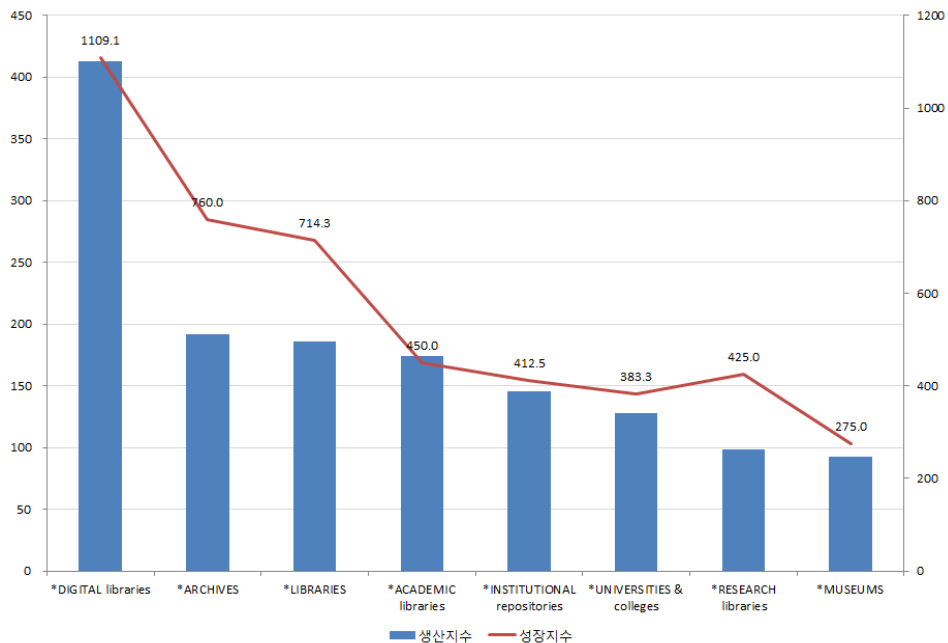


서비스(서비스 중심)로 전환되는 과정에 있음을 알 수 있다. 특히 출판(데이터), 큐레이터직, 교육과정의 성장이 두드러지고 연구, 데이터 분석, 사서, 교육, 아키비스트, 정보전문가의 성장 추세가 뚜렷하여 전통적인 정보전문직 교육과정의 변화를 통한 큐레이터 양성 및 데이터 서비스의 발전 가능성이 큰 것으로 나타났다. 따라서 차세대 정보서비스로서 디지털 큐레이션의 지향점이 되는 이 영역의 향후 연구 과제는 크게 세 가지로 구분할 수 있다. 첫째, 큐레이터 및 이용자에 대한 디지털 큐레이션 교육 및 훈련 프로그램 개발 및 제공. 둘째, 연구 측면에서 개인이나 소규모 집단이 수행하는 연구에서 생산되어 개인 컴퓨터에서 사용 및 조작성이 가능한 소규모(또는 중규모) 데이터에 대한 큐레이팅과 함께 연구성과물(학술논문, 단행본 등)이 생

산되는 과정에서 수집·가공·생산된 데이터 출판의 지원. 셋째, 데이터를 포함한 디지털 문화유산의 재사용에 중점을 둔 부가서비스(기획 및 분석 등) 개발 및 제공.

### 4.3 기관

〈그림 11〉은 기관 영역의 생산성과 성장성을 나타낸 것이다. 생산지수와 성장지수 양자에서 가장 높은 디지털 도서관을 비롯하여 도서관, 기록관, 기관 리포지토리, 박물관 등 주요 문화유산기관들이 모두 디지털 큐레이션 관련 기관에 포함되어 있다. 특히, 최근에는 기록관, 대학 및 연구도서관의 성장이 두드러지게 나타나고 있어 이들 기관을 중심으로 연구가 이루어지고 있음을 알 수 있다.



〈그림 11〉 기관 영역의 주제별 생산지수와 성장지수

기관 영역의 우선적인 연구과제로는 기록관과 연구기관(대학, 연구소 등)의 활동과 연계한 디지털 큐레이션에 중점을 둘 필요가 있다. 중·장기적으로는 현재까지 저조하거나 불모지로 남아있는 박물관과 공공도서관 측면에서의 디지털 큐레이션을 모색하는 연구가 필요할 것이다. 또한, 새롭게 출현한 문화유산기관으로서 리포지토리(주제 리포지토리, 기관 리포지토리)를 통한 디지털 큐레이션 관련 연구도 필요하며, 나아가 이들 주요 문화유산기관의 디지털 큐레이션을 통합 및 연계하기 위한 연구도 이루어져야 한다.

## 5. 결 론

디지털 형식으로 생산 및 변환된 모든 정보는 디지털 큐레이션의 대상이 될 수 있으며, 인식 가능한 디지털 정보 객체는 가치 있는 인류의 문화유산으로서 차세대 정보관리 및 서비스의 기본 단위가 된다. 또한, 아날로그 환경에서 대표적인 정보관리기관이었던 도서관, 기록관, 박물관 등은 물론 새롭게 등장한 리포지토리가 디지털 환경의 대표적인 문화유산기관으로 디지털 큐레이션의 핵심 주체가 되고 있다. 사전에 정보관리가 선행되어야만 정보서비스가 가능하다는 측면에서, 각 기관들의 디지털 정보관리를 중심으로 시작되어 급속히 성장하고 있는 디지털 큐레이션은 차세대 정보서비스로서 그 범위와 방법이 확장 및 고도화되고 있다. 초기에는 단순히 기존의 아날로그 장서관리의 일부로 시작되었으나 점차 독립적인 디지털 정보관리 영역으로 성장하는 동시에 정보관리 및 서비스의 대

상으로서 데이터가 중요한 의미를 갖게 되었다.

본 연구는 문헌정보학 분야를 중심으로 디지털 큐레이션에 관한 연구 동향을 여러 측면에서 분석한 결과에 기초하여 차세대 디지털 정보관리 및 서비스의 주체로서 위상을 확립하기 위한 향후의 연구 과제를 제시하였다. 이를 위해 디지털 큐레이션 관련 문헌들을 검토하여 연구 영역을 설정하고, 문헌정보학 분야의 대표적인 학술 데이터베이스(LISTA)에서 검색한 학술논문을 대상으로 다양한 분석 기법을 적용하여 연구 동향을 분석한 결과에 기초하여 향후의 연구 과제를 제시하였다. 먼저 디지털 큐레이션 관련 문헌들을 검토한 결과로서 연구 영역은 크게 이론 및 인프라, 정보관리 및 서비스, 기관의 3개 영역과 이들의 하위 주제들로 설정하였다. 다음으로 LISTA에서 검색한 374편의 학술논문을 대상으로 기본 분석과 지적구조 분석을 수행하여 실제적인 연구 동향을 파악하였다. 문헌정보학 분야 디지털 큐레이션 연구는 2000년대 이후 본격적으로 시작되어 최근 급속히 성장하고 있다. 지적구조 분석을 통해 시기별로 살펴본 결과로는 1기(1987년~2000년)에 개별 기관의 정보관리 및 서비스에 정보기술을 일부 적용하는 차원의 연구에서, 2기(2001년~2007년)에 주요 문화유산기관들의 디지털 정보관리 및 서비스를 중심으로 하는 디지털 큐레이션 관련 연구가 본격적으로 시작되었다. 최근의 3기(2008년~2014년)는 이전의 정보관리 영역의 연구가 주종을 이루는 가운데 서비스로의 전환이 시도되고 있다. 또한, 영역별로는 9개 군집(영역)을 형성하면서 디지털 도서관, 정보자원 관리, 디지털 보존을 중심으로 연구가 진행되었다. 또한, 생산성과 성장성 측면을 종합적으로

분석한 결과에서는 디지털 큐레이션의 대상이 되는 정보자원의 개념과 실행 주체로서 기관의 확장 추세가 뚜렷하게 나타났다. 즉, 문헌정보학 분야의 디지털 큐레이션 연구는 이전에는 개별 기관의 유형별 정보관리가 주류를 이루고 있으나, 최근에 디지털 큐레이션의 대상(문화유산, 데이터) 및 주체(문화유산기관: 도서관, 기록관, 박물관, 리포지토리)의 확장 추세와 함께 그 중심이 점진적으로 정보관리에서 서비스로 전환되는 과정에 있다고 할 수 있다.

연구 동향을 분석한 결과를 검토하여 디지털 큐레이션의 연구 영역별로 도출한 향후 연구 과제는 다음과 같다. 첫째, 이론 및 인프라 영역의 향후 연구 과제로는 국가적 인프라(e-Research, e-Science, cyberinfrastructure, platform 등)의 구축과 연계하여 효율적인 디지털 큐레이션을 지원하기 위한 이론과 모형, 정책 및 전략에 관한 연구에 중점을 둘 필요가 있다. 특히, 이론 영역에서 각 기관의 장서관리를 넘어 문화유산 기관들의 디지털 문화유산을 대상으로 새로운 정보관리 및 서비스를 추진하기 위한 디지털 큐레이션의 이론적 토대와 전략을 마련하는 연구는 시급하다.

둘째, 정보관리 및 서비스 영역의 연구 과제로는 정보관리 측면에서 디지털 큐레이션의 대상이 되는 정보자원으로서 디지털 문화유산의 전반적인 관리에 관한 연구와 함께 새로운 정보자원인 데이터 관리에 관한 연구가 중점적으로 이루어져야 할 것이다. 또한 일부 주제(보존, 디지털화 등)에 치중된 경향을 벗어나 가치 있는 디지털 문화유산의 선정 및 수집, 재사용, 아카이빙, 권한 관리 및 보안, 메타데이터, 파일 포맷 및 인코딩 등 다양하고 실질적인 주제들

을 다루어야 한다. 이외에 현재와 같이 관리 지향의 보존 및 수집 중심이 아닌 서비스 측면의 아카이빙과 큐레이팅을 위한 연구가 필요하며, 이와 같은 맥락에서 그 토대가 되는 메타데이터에 관한 연구도 활성화되어야 한다. 디지털 큐레이션의 지향점이라 할 수 있는 정보서비스 측면의 향후 연구 과제는 세 가지로 구분할 수 있다. 첫째, 큐레이터 및 이용자에 대한 디지털 큐레이션 교육 및 훈련 프로그램 개발. 둘째, 연구 측면에서 개인이나 소규모 집단이 수행하는 연구에서 생산되어 개인 컴퓨터에서 사용 및 조작이 가능한 소규모(또는 중규모) 데이터에 대한 큐레이팅과 함께 연구성과물(학술논문, 단행본 등)이 생산되는 과정에서 수집·가공·생산된 데이터 출판의 지원. 셋째, 데이터를 포함한 디지털 문화유산의 재사용에 중점을 둔 부가서비스(기획 및 분석 등) 개발 및 제공.

셋째, 기관 영역의 우선적인 연구과제로는 기록관과 교육 및 연구기관(대학, 연구소 등)의 활동과 연계한 디지털 큐레이션에 중점을 둘 필요가 있다. 중·장기적으로는 현재까지 저조하거나 불모지로 남아있는 박물관과 공공도서관 측면에서의 디지털 큐레이션을 모색하는 연구가 필요할 것이다. 또한, 새롭게 출현한 문화유산기관으로서 리포지토리(주제 리포지토리, 기관 리포지토리)를 통한 디지털 큐레이션 관련 연구도 필요하며, 나아가 이들 주요 문화유산기관의 디지털 큐레이션을 통합 및 연계하기 위한 연구도 이루어져야 한다.

본 연구의 제한점으로 첫째, 전체 학문 분야를 대상으로 하지 않고 문헌정보학 분야를 중심으로 국외의 연구동향을 살펴보았다. 둘째, 분석 대상이 되는 문헌집합을 학술논문으로 제한

하여 학위논문, 단행본 등은 포함하지 않았다. 셋째, 학술지 논문에 부여된 디스크립터를 중심으로 분석하였고, 다른 관점에서 연구동향을 파악할 수 있는 연구자 또는 연구집단에 대한 분석은 수행하지 않았다. 따라서 본 연구의 결과

를 토대로 향후 문헌 유형을 확장한 문헌집합에 기초한 종합적인 분석, 그리고 또 다른 관점에서 연구동향을 파악할 수 있는 것으로 연구주체(연구자 또는 연구집단)에 관한 분석 등의 후속 연구가 지속적으로 필요할 것이다.

## 참 고 문 헌

- 김지현 (2014). 대학도서관의 연구데이터관리서비스에 관한 연구 - 미국 연구중심대학도서관을 중심으로. 한국비블리아학회지, 25(3), 165-189. <http://dx.doi.org/10.14699/kbiblia.2014.25.3.165>
- 김관준, 이재윤 (2007). 연구 영역 분석을 위한 디스크립터 프로파일링에 관한 연구. 정보관리학회지, 24(4), 285-303. <http://dx.doi.org/10.3743/KOSIM.2007.24.4.285>
- 신영란, 정연경 (2012). 국내 인문사회 연구데이터 아카이브의 개선방안에 관한 연구. 한국기록관리학회지, 12(3), 93-115.
- 안영희, 박옥화 (2010). 디지털 큐레이션 정책을 위한 프레임워크 개발. 한국도서관·정보학회지, 41(1), 167-186.
- 이소연 (2012). 함께 만드는 미래: 디지털 융합과 문화유산기관의 협력. 정보관리학회지, 29(3), 235-256.
- 이재윤 (2006). 계량서지적 네트워크 분석을 위한 중심성 척도에 관한 연구. 한국문헌정보학회지, 40(3), 191-214. <http://dx.doi.org/10.4275/KSLIS.2006.40.3.191>
- 이재윤 (2013). tnet과 WNET의 가중 네트워크 중심성 지수 비교 연구. 정보관리학회지, 30(4), 241-264.
- 한국기록학회 편 (2008). 기록학 용어사전. 2008. 서울: (주)역사비평사.
- Abbott, D. (2008). What is digital curation? [cited 2013.10.1] Retrieved from <http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/what-digital-curation>
- Alexandria Digital Library (2014). Curation bibliography. [cited 2013.11.17] Retrieved from <http://legacy.alexandria.ucsb.edu/~gjanee/dc@ucsb/bibliography.html>
- Antell, K., Foote, J. B., Turner, J., & Shults, B. (2014). Dealing with data: Science librarians' participation in data management at Association of Research Libraries Institutions. College & Research Libraries, 75(4), 557-574. <http://dx.doi.org/10.5860/crl.75.4.557>
- Bailey, Charles, W., Jr. (2012). Digital duration bibliography: Preservation and stewardship of scholarly works. Houston, TX: Digital Scholarship. [cited 2013.11.3] Retrieved from <http://digital-scholarship.org/dcbw/dcb.pdf>

- Bailey, Charles, W., Jr. (2014). Research data curation bibliography, version 4. [cited 2014.1.27]  
Retrieved from <http://digital-scholarship.org/rdcb/rdcb.htm>
- Carlson, S. 2006. Lost in a sea of science data. *Chronicle of Higher Education*, 52(42). [cited 2013.12.1]  
Retrieved from [http://sdl.syr.edu/course/week1/1\\_IntroductionToTheCourse\\_Carlson\\_2006.pdf](http://sdl.syr.edu/course/week1/1_IntroductionToTheCourse_Carlson_2006.pdf)
- Choudhury, G. S. (2008). Case study in data curation at Johns Hopkins University. *Library Trends*, 57(2), 211-220.
- Choudhury, G. S., & Stinson, T. L. (2007). The virtual observatory and the roman de la rose: Unexpected relationships and the collaborative imperative. *Academic Commons, Special Cyberinfrastructure Issue*. [cited 2013.11.21] Retrieved from  
<http://www.academiccommons.org/commons/essay/VO-and-roman-de-la-rose-collaborative-imperative>
- Choudhury, G. S., and Lynch, C. (2009). Initiatives from the NSF's DataNet program: DataONE and the data conservancy. [cited 2013.12.15] Retrieved from  
<http://www.educause.edu/annual-conference/2009/initiatives-nsfs-datanet-program-dat-aone-and-data-conservancy-sponsored-juniper-networks-qwest-communi>
- Digital Curation Center (2013). What is digital curation? [cited 2013.11.1] Retrieved from  
<http://www.dcc.ac.uk/digital-curation/what-digital-curation>
- Elliot, Gillian (2009). Biodiversity data management project: Extending the boundaries of information management in collaboration with life scientists at the University of Otago. *New Zealand Library & Information Management Journal*, 51(2), 104-120.
- Fearon, D. Jr., Gunia, B., Pralle, E., Lake, S., & Sallans, A. L. (2013). SPEC Kit 334: Research data management services. Washington, DC: ARL.
- Gold, A. (2007). Cyberinfrastructure, data, and libraries, Part 2: Libraries and the data challenge: Roles and actions for libraries, *D-Lib Magazine*, 13(9/10). [cited 2013.10.29] Retrieved from <http://www.dlib.org/dlib/september07/gold/09gold-pt2.html>
- Gold, A. (2010). Data curation and libraries: Short-term developments, Long-Term Prospects. [cited 2014.1.13] Retrieved from [http://digitalcommons.calpoly.edu/lib\\_dean/27/](http://digitalcommons.calpoly.edu/lib_dean/27/)
- Higgins, S. (2008). The DCC curation lifecycle model. *The International Journal of Digital Curation*, 1(3), 134-140. [cited 2013.11.15] Retrieved from  
<http://www.ijdc.net/index.php/ijdc/article/viewFile/69/48>
- Higgins, S. (2011). Digital curation: The emergence of a new discipline. *The International Journal of Digital Curation*, 6(2), 78-88. [cited 2014.12.13] Retrieved from  
<http://www.ijdc.net/index.php/ijdc/article/view/184>

- Jahnke, Lori M., & Asher, Andrew (2012). The problem of data: Data management and curation practices among university researchers, CLIR pub154. [cited 2013.12.13] Retrieved from <http://www.clir.org/pubs/reports/pub154/problem-of-data>
- Joint Task Force on Library Support for E-Science (2007). Agenda for developing E-science in research libraries. Washington, DC: Association of Research Libraries. [cited 2013.11.13] Retrieved from [http://www.arl.org/bm~doc/ARL\\_EScience\\_final.pdf](http://www.arl.org/bm~doc/ARL_EScience_final.pdf)
- Jones, S., Pryor, G., & Whyte, A. (2013). How to develop research data management services – A guide for HEIs. Edinburgh: Digital Curation Centre.
- Lee, Christopher A., & Tibbo, Helen R. (2007). Digital curation and trusted repositories: Steps toward success. *Journal of Digital Information*, 8(2). [cited 2013.11.13] Retrieved from <https://journals.tdl.org/jodi/index.php/jodi/article/view/229/183>
- Lynch, C. (2006). Open computation: Beyond human-reader-centric views of scholarly literature. In *Open access: Key strategic, technical and economic aspects.*, ed. Neil Jacobs, 185-193. Oxford: Chandos.
- McCue, J., & Corson-Rikert, J. (2007). Engaging and connecting faculty: Research discovery, access, re-use, and archiving. Project Briefing, CNI. [cited 2013.10.13] Retrieved from <http://www.cni.org/tfms/2007a.spring/abstracts/PB-engagingmccue.html>
- Messerschmitt, D. G. (2003). Opportunities for research libraries in the NSF cyberinfrastructure program. *ARL Bimonthly Report*, 229, 1-7. [cited 2014.11.15] Retrieved from <http://old.arl.org/resources/pubs/br/br229/br229cyber.shtml>
- NSB (National Science Board) (2005). Long-lived digital data collections: Enabling research and education in the 21st century. *NSF*, 37. [cited 2014.11.17] Retrieved from <http://www.nsf.gov/pubs/2005/nsb0540>
- Purdue University Libraries (2007). D2C2 - Purdue University. [cited 2014.11.11] Retrieved from <http://d2c2.lib.purdue.edu/>
- Rusbridge, C., et al. (2005). The digital curation centre: A vision for digital curation. In: *Proceedings of local to global data interoperability. Challenges and Technologies*, 20-24 June 2005, Sardinia, Italy. [cited 2014.11.13] Retrieved from <http://eprints.erpanet.org/82/>
- Simons, N., & Richardson, J. (2013). *New content in digital repositories: The changing research landscape*. Chandos Information Professional Series, Oxford Cambridge New Delhi: CP.
- Steinhart, G., Dietrich, D., & Green, A. (2009). Establishing trust in a chain of preservation: The TRAC checklist applied to a data staging repository (DataStaR). *D-Lib Magazine* 15(9/10). [cited 2014.12.21] Retrieved from

<http://www.dlib.org/dlib/september09/steinhart/09steinhart.html>

Steinhart, G., Saylor, J., & Westbrook, E. L. (2008). Digital research data curation: Overview of issues, current activities, and opportunities for the Cornell University Library.

[cited 2014.12.21] Retrieved from <http://ecommons.cornell.edu/handle/1813/10903>.

Witt, M. (2012). Co-designing, Co-developing, and co-implementing an institutional data repository service. *Journal of Library Administration*, 52(2), 172-188.

Wolski, Malcolm and Richardson, J. (2010). Moving researchers across the eresearch chasm. *Ariadne*, 65. Bath, UK: UKOLN. [cited 2014.12.28] Retrieved from

<http://www.ariadne.ac.uk/issue65/wolski-richardson/>

• 국문 참고문헌에 대한 영문 표기

(English translation of references written in Korean)

Ahn, Young-Hee, & Park, Ok-Wha (2010). Development of a framework for digital curation policy. *Journal of Korea Library and Information Science Society*, 41(1), 167-186.

Kim, Jihyun (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. <http://dx.doi.org/10.14699/kbiblia.2014.25.3.165>

Kim, Pan Jun, & Lee, Jae Yun (2007). Descriptor profiling for research domain analysis. *Journal of the Korean Society for Information Management*, 24(4), 285-303. <http://dx.doi.org/10.3743/KOSIM.2007.24.4.285>

Lee, Jae Yun (2006). Centrality measures for bibliometric network analysis. *Journal of the Korean Society for Library and Information Science*, 40(3), 191-214. <http://dx.doi.org/10.4275/KSLIS.2006.40.3.191>

Lee, Jae Yun (2013). A comparison study on the weighted network centrality measures of tnet and WNET. *Journal of the Korean Society for Information Management*, 30(4), 241-264.

Lee, So-Yeon (2012). The future we make together: Cultural heritage institutions in the era of digital convergence. *Journal of the Korean Society for Information Management*, 29(3), 235-256.

Records Society of Korea (2008). *Glossary of archives*. Seoul: History Criticism, Co.

Shin, Young-Ran, & Chung, Yeon-Kyoung (2012). A study on the improvement plans of the humanities and social sciences research data archives in Korea. *Journal of Records Management & Archives Society of Korea*, 12(3), 93-115.