

# 도서관 빅데이터 서비스 모형 개발에 관한 연구: 공공도서관을 중심으로\*

## A Study on the Developing of Big Data Services in Public Library

표순희 (Soon Hee Pyo)\*\*

김윤형 (Yun Hyung Kim)\*\*\*

김혜선 (Hye Sun Kim)\*\*\*\*

김완종 (Wan Jong Kim)\*\*\*\*\*

### 초 록

본 연구는 최근 많은 이슈가 되고 있는 빅데이터를 도서관 분야에 적용하여 다양한 형태의 도서관 빅데이터의 활용 가치에 대한 이해를 높이고 이에 대한 수요자의 요구 분석을 바탕으로 공공도서관 빅데이터 서비스 모형을 개발하는 것을 목적으로 하고 있다. 이를 위해 도서관 빅데이터의 개념과 내용 및 가치 등을 고찰하고, 도서관 빅데이터 서비스에 대한 수요 분석을 바탕으로 도서관 빅데이터 서비스 모형을 개발하였다. 서비스 모형 개발을 위해 도서관 빅데이터의 유형에 따라 활용 가능한 도서관 빅데이터를 분석하였으며, 수요자의 요구를 다양한 방법으로 도출하였다. 수요자의 요구 분석은 도서관계 연구자 및 현장 사서와의 심층인터뷰, 표적집단인터뷰(Focus Group Interview, 이하 FGI), 사서 및 이용자 설문조사를 통해 이루어졌다. 이를 바탕으로 총 16개의 도서관 빅데이터 서비스 모형을 정의하고, 서비스의 필요성, 시급성, 개발 가능성을 고려해 최종적으로 사서 의사결정 지원 서비스와 이용자 도서 추천 및 독서이력 관리 서비스 모형을 개발하였다.

### ABSTRACT

Big data refers to dataset whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze. And now it is considered to create the new opportunity in every industry. The purpose of this study is to develop of big data services in public library for improved library services. To this end, analysed the type of library big data and needs of stockholders through the various methods such as deep interview, focus group interview, questionnaire. At first step, we defined the 16 big data service models from interview with librarians, and LIS professions. Second step, it was considered necessity, timeliness, possibility of development. We developed the final two services called on 'Decision Support Services for Public Librarians' and 'Book Recommendation Services for Users.'

키워드: 공공도서관, 빅데이터, 빅데이터 서비스

public library, big data, big data service, big data value

\* 본 연구는 2014년도 문화체육관광부의 지원을 받아 연구되었음.

\*\* 성균관대학교 정보관리연구소 선임연구원(shpyo88@hanmail.net) (제1저자)

\*\*\* (주)기술과가치 선임연구원(yhkim@technovalue.com) (제2저자)

\*\*\*\* 한국과학기술정보연구원 정보서비스실 실장(hskim@kisti.re.kr) (제3저자)

\*\*\*\*\* 한국과학기술정보연구원 정보서비스실 선임연구원(wjkim@kisti.re.kr) (교신저자)

■ 논문접수일자: 2015년 5월 27일 ■ 최초심사일자: 2015년 5월 27일 ■ 게재확정일자: 2015년 6월 12일

■ 정보관리학회지, 32(2), 63-86, 2015. [<http://dx.doi.org/10.3743/KOSIM.2015.32.2.063>]

## 1. 서론

### 1.1 연구의 필요성 및 목적

정보통신 또는 스마트 혁명과 함께 사회의 다양한 곳에서 많은 사람들이 데이터 양이 폭증하고 데이터 유형이 다양화되면서, 이들 데이터를 수집, 축적, 분석, 활용하여 새로운 가치를 만들어 내는 빅데이터가 정보통신 분야 및 산업 전 분야에 새로운 이슈로 부상하고 있다. 특히 미국, 유럽, 일본 등 해외 주요국은 공공정보 이용을 촉진하거나 공공부문이 직접 빅데이터 기술개발을 지원하는 방식으로 성장세에 있는 민간부분을 지원하고, 다른 한편으로 개인 정보보호, 인프라 구축 지원 등 제도적 기반을 개선하는 방식으로 빅데이터 진흥 정책을 추진하고 있다(배동민 외, 2013).

우리나라도 빅데이터의 중요성을 인식하여 공공데이터의 공개 및 스마트 국가 구현을 위한 빅데이터 마스터 플랜 구축 등 정부차원에서 지속적으로 노력하고 있다(관계부처 합동, 2012). 이에 빅데이터를 활용한 연구자 및 산업체의 연구 개발이 증가하면서 빅데이터 수요 급증에 따른 빅데이터 서비스가 확대되고 있다. 정보기술의 발달로 인해 모바일 서비스 환경에서 스마트 기기 등을 통해 언제 어디서든지 독서를 즐길 수 있는 도서관 환경의 변화와 함께 이러한 이용자들에 의해 생산되고 축적된 도서관 관련 데이터는 도서관 빅데이터 서비스에 대한 수요를 증가시키는 자극제가 되고 있다. 도서관분야에서도 2014년 '제2차 도서관종합발전계획'을 발표하며, 빅데이터를 활용한 서비스를 개발하고자 하는 의지를 표명하고 있다

(도서관정보정책위원회, 2014).

이러한 시점에서 본 연구는 도서관 빅데이터 및 데이터의 활용 가치에 대한 이해를 높이고 도서관 빅데이터 수요자의 요구 분석을 바탕으로 하여 공공도서관 빅데이터 서비스 모형을 개발하여 변화하는 환경에 맞는 향상된 도서관 서비스를 제공하고자 한다.

### 1.2 연구 내용 및 방법

본 논문의 연구 내용은 크게 3가지로 구분할 수 있다. 첫째, 도서관 빅데이터의 개념과 내용을 고찰하였다. 빅데이터라는 개념이 최근 정보의 폭발적인 증가와 축적으로 다양한 분야에서 발생한 것으로 도서관 측면에서 빅데이터의 개념, 내용, 특성 및 분류체계에 대해 살펴봄으로써 도서관 빅데이터에 대한 이해를 돕고자 하였다. 둘째, 서비스 모형 개발을 위해 도서관 빅데이터의 유형에 따라 활용 가능한 도서관 빅데이터를 분석하였으며, 수요자의 요구를 다양한 방법으로 도출하였다. 셋째, 이를 바탕으로 공공도서관에 필요한 최종적인 도서관 빅데이터 서비스 모형을 도출하고, 이에 필요한 개발방안을 제시하였다.

빅데이터 서비스 모형은 문헌 연구, 심층 인터뷰, 전문가를 대상으로 한 FGI, 설문조사와 같이 다양한 연구 방법으로 이루어졌다. 문헌 연구는 도서관 빅데이터의 정의와 개념, 내용, 특성 등 도서관 빅데이터에 대한 이해를 높이기 위해 수행되었다. 또한, 국내의 도서관 분야의 빅데이터 활용사례를 조사, 분석하고 도서관 빅데이터 서비스의 성과라는 관점에서 빅데이터의 가치를 고찰하였다. 심층 인터뷰, 전문

가를 대상으로 한 FGI, 사서 및 이용자 설문조사와 같은 조사연구는 도서관 빅데이터 서비스의 핵심고객을 규명, 도서관 업무별 애로사항 및 개선 요구사항을 조사, 분석하고, 도서관에서 요구되는 구체적인 빅데이터 서비스 수요를 도출하기 위해 이루어졌다.

본 연구는 최근 급격히 증가하는 도서관의 내외부 데이터의 통합 및 분석의 필요성에 대한 빅데이터 서비스 개발을 위한 것으로 서비스 수요에 따른 서비스 모형 개발에 초점을 두었다. 따라서 빅데이터 서비스와 시스템 인프라, 분석 플랫폼 및 서비스 지원 기술과 관련된 기술 아키텍처 부분은 본 연구에 포함하지 않았다.

## 2. 도서관의 빅데이터와 서비스

### 2.1 빅데이터의 정의와 특성

빅데이터의 특징에 대해 아직까지 단일한 개념이 정립되어 있지 않으나 가장 일반적으로 데이터의 크기(Volume), 다양한 형태(Variety), 빠른 생성·유통·속도(Velocity)라는 뜻에서 3Vs라는 특징을 가장 많이 사용하고 있다. 최근에는 3Vs에 빅데이터를 통한 가치(Value) 창출의 중요성이 강조되면서 4Vs, 4Vs에 다시 데이터의 진본성(Veracity)을 더하여 5Vs라 정의하기도 한다.

빅데이터의 개념에 대한 정의를 살펴보면 Gartner(2012)는 빅데이터를 '향상된 시사점을 제공하고, 더 나은 의사결정을 위해 사용되는 데이터 가운데 비용 효율이 높고, 혁신적이며,

대용량, 고속 및 다양성의 특성을 가진 정보 자산' 이라고 정의하였고 Mckinsey Global Institute (2011)는 데이터베이스의 규모에 초점을 맞추어, '일반적인 데이터베이스 소프트웨어가 저장, 관리, 분석할 수 있는 범위를 초과하는 규모의 데이터'라고 정의하였다. IDC(2011)는 데이터베이스가 아닌 업무수행에 초점을 맞추어, '다양한 종류의 대규모 데이터로부터 저렴한 비용으로 가치를 추출하고 데이터의 초고속 수집, 발굴, 분석을 지원하도록 고안된 차세대 기술 및 아키텍처'로 개념화하였다. 빅데이터는 활용 목적에 따라 다양하게 정의하고 있지만 요약하면 '기존의 방식으로는 저장, 관리, 분석이 어려울 정도로 규모가 크고 데이터의 순환, 처리 등의 속도가 빠르며, 형식이 다양한 데이터 또는 이러한 데이터를 분석하는 방법'을 통칭한다.

이러한 빅데이터의 개념을 외부 데이터, 비정형, 실시간 데이터와 이형의 데이터 결합으로 새로운 지식을 창출하는 새로운 영역의 데이터로 그 의미를 확장하기도 한다(최성근 외, 2012). 김서(2013)도 빅데이터를 구성하는 데이터 종류 및 속성을 정형데이터, 반정형데이터, 비정형데이터로 구분하여 빅데이터의 데이터 유형을 구분하여 실제 생활에서 생성되는 빅데이터의 방대함과 활용성에 대한 이해를 높였다. 홍승필 외(2012)는 <표 1>과 같이 빅데이터 요소기술을 데이터 수집/통합, 데이터 저장/관리, 데이터 분석, 데이터 분석 가시화라는 네 가지로 구분하여 설명하고 있으며 자세한 내용은 다음과 같다.

먼저, 데이터 수집은 분산된 다양한 자원으로부터 필요로 하는 데이터를 수동 혹은 자동

〈표 1〉 빅데이터 요소기술 분석

분류	정의	빅데이터 요소기술	
		관련 기술	오픈소스 기술
빅데이터 수집/통합	새로운 데이터 생성, 네트워크에 산재해 있는 외부 데이터 수집, 내·외부 이종데이터 통합 등 데이터의 형태와 소재에 무관하게 데이터를 확보하는 기술	<ul style="list-style-type: none"> <li>• 에이전트</li> <li>• 웹로봇</li> <li>• RSS</li> <li>• 웹 크롤러</li> </ul>	<ul style="list-style-type: none"> <li>• Scribe</li> <li>• Flume</li> <li>• Chukwa</li> </ul>
데이터 저장/관리	폭증하는 다양한 형식의 데이터를 실시간 저장/관리할 수 있는 분산 컴퓨팅기술	<ul style="list-style-type: none"> <li>• 대용량 분산 파일 시스템</li> <li>• 인-DB 분석</li> <li>• 인-메모리 기술</li> </ul>	<ul style="list-style-type: none"> <li>• NoSQL</li> <li>• 빅테이블(Big Table)</li> <li>• GFS(구글 파일시스템)</li> </ul>
데이터 분석	빅데이터에 내재된 가치를 추출하기 위해 필요한 대규모 통계처리, 데이터 마이닝, 그래프 마이닝 등의 분석 방법, 기계학습 및 인공지능을 활용한 심층 분석 기술	<ul style="list-style-type: none"> <li>• Text Mining</li> <li>• 자연어 처리</li> <li>• Machine Learning</li> <li>• DDS</li> <li>• RDBMS</li> </ul>	<ul style="list-style-type: none"> <li>• MapReduce</li> <li>• Big Query</li> <li>• Mahout</li> </ul>
데이터 분석 가시화	비전문가가 데이터 분석을 수행할 수 있는 환경을 제공하는 분석 도구 기술과 분석 결과를 함축적으로 표시하고, 직관적인 정보를 제공하는 인포그래픽스 기술로 구성	<ul style="list-style-type: none"> <li>• InVis</li> <li>• Spatial Information Flow</li> <li>• Clustergram</li> <li>• History Flow</li> <li>• Facebook Transaction</li> </ul>	<ul style="list-style-type: none"> <li>• R</li> </ul>

으로 수집하는 과정으로, 조직 내부에 분산된 정형 데이터의 수집과 조직 외부에 흩어진 비정형 데이터의 수집을 모두 고려한다. 둘째, 빅데이터는 기존 정형화된 데이터뿐만 아니라 비정형화된 데이터를 포함하고, 규모 또한 방대하여 기존의 데이터베이스로는 저장 및 관리에 한계가 있기 때문에 새로운 형태의 데이터베이스 및 DBMS가 등장하고 있다. 셋째, 빅데이터를 분석하기 위한 기법들은 통계학과 전산학, 특히 기계학습과 데이터마이닝 분야에서 이미 사용된 기법들이다. 이 분석기법들의 알고리즘을 대규모 처리에 맞도록 개선하여 빅데이터 처리에 적용하는 기술로 텍스트/오피니언 마이닝, 소셜네트워크분석, 군집분석 등이 주목을 받고 있다. 마지막으로, 빅데이터 분석 가시화 기술은 비전문가가 데이터 분석을 수행할 수 있는 환경을 제공하는 분석도구 기술과 분석결

과를 함축적으로 표시하고 직관적인 정보를 제공하는 R, InVIS 등 인포그래픽스 기술로 구성된다.

## 2.2 도서관 빅데이터와 특성

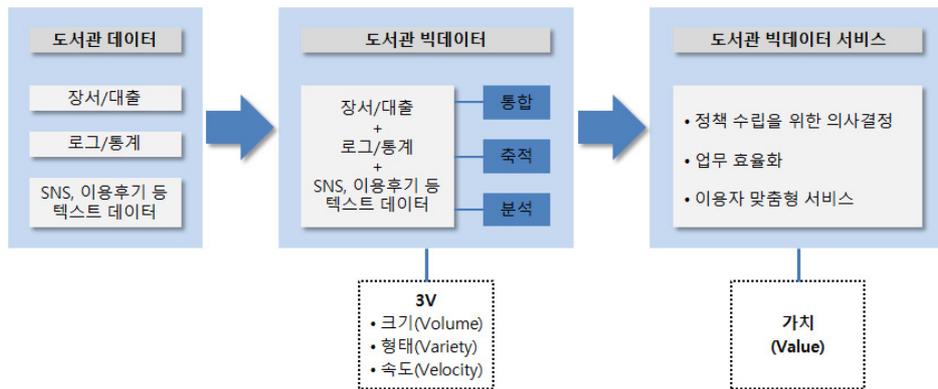
도서관의 데이터는 도서관 내외부에서 수집, 생성되는 다양한 데이터이다. 도서관의 빅데이터는 이러한 데이터 중 빅데이터의 특징인 3Vs를 만족하고, 빅데이터의 다양성 측면에서 논할 수 있는 3가지 유형인 정형, 반정형, 비정형 데이터에 속하는 데이터를 의미한다. 따라서 도서관 빅데이터는 도서관 정보와 빅데이터의 특성을 합성하여 '도서관 내외부에서 수집, 생성되는 다양한 도서관 정보로 빅데이터의 특징(3Vs)과 유형에 부합하고 도서관 분야의 가치를 창출할 수 있는 데이터'로 정의할 수 있다.

즉, 도서관 빅데이터는 정형, 반정형, 비정형 형태의 도서관 정보가 빠른 속도로 생산되고 방대하게 통합, 축적되어 도서관 정책 실무자, 도서관 사서, 도서관 이용자를 위해 분석되고 서비스를 개발하는데 활용되는 데이터로 <그림 1>과 같이 도식화 할 수 있다.

이러한 도서관 데이터는 빅데이터의 다양성 측면에서 바라 본 3가지 유형에 따라 다음과 같이 분류할 수 있다.

첫째, 도서관 정형데이터는 도서관 홈페이지 및 도서관 정보시스템을 통해 수집되는 정보로 관계형 데이터베이스 형태의 DBMS에 저장되는 도서관 데이터이다. 도서관 홈페이지 웹로

그 데이터, 도서관 이용자 검색로그 데이터, 수서시스템에 기록된 수서 데이터, 통합서지용 KORMARC로 기록된 목록데이터 등의 도서관 장서데이터, 도서 대출·반납 정보, 열람데이터 등 도서관 이용데이터 등이 정형데이터에 해당한다. 둘째, 도서관 반정형데이터는 도서관에서 수집되는 정보를 바탕으로 도서관에서 자체적으로 생성하는 데이터로, 가공 기술에 따라 다양한 데이터로 변형, 분석이 가능한 데이터이다. 대표적으로 시설 및 대출 현황 등의 도서관 통계 데이터와 공공데이터포털(data.go.kr)을 통해 제공되는 도서관 분야 Open API, 데이터셋 자료 등이 해당한다. 마지막으로, 도서관 비정형



<그림 1> 도서관 빅데이터와 서비스의 개념도

<표 2> 다양성 측면에서의 도서관 빅데이터 분류체계

구분	내용	종류
도서관 정형데이터	도서관 홈페이지 및 정보시스템을 통해 수집되는 정보들로, 관계형 데이터베이스 형태의 DBMS에 저장되는 전형적인 도서관 데이터	장서데이터 이용데이터
도서관 반정형데이터	도서관에서 수집되는 정보를 바탕으로 도서관에서 자체적으로 생성하는 데이터	로그데이터 통계데이터 공공데이터
도서관 비정형데이터	게시판, SNS 등을 통해 생성되는 텍스트정보	도서관 관련 텍스트 정보

데이터는 도서관 내/외부에서 생성되는 도서관 사서 및 도서관 이용자들이 생성하는 텍스트 정보이다. 도서관 홈페이지 게시판 및 SNS 등에 게시되는 텍스트 정보 등의 도서관 내부 데이터와 지역별, 목적별로 도서관 사서 및 도서관 이용자들이 구축, 운영하고 있는 홈페이지 및 블로그 등에 생산되는 도서관 관련 텍스트 정보, 도서관 문화프로그램 이용자들의 이용후기 데이터 등이 비정형데이터에 해당 한다(〈표 2〉 참조).

### 2.3 국내외 도서관 빅데이터 활용

도서관에서 빅데이터의 활용은 정보의 디지털 아카이빙에서 출발한다. 미국 의회도서관은 단문기록을 위해 트위터(twitter.com)에 게시된 모든 글들을 영구보존하는 프로젝트를 진행 중에 있다. 미국 의회도서관은 트위터와의 계약을 통해 타임라인에 게시되는 모든 트윗에 대한 접근권한을 보유하게 되었으며, 2012년 12월 기준 1,700억 건의 트윗을 포함한 총 133 테라바이트 규모의 아카이브를 보유하고 있다. 현재 이 아카이브에 대한 공개적인 접근을 제공하고 있지 않지만, SNS를 통해 생산된 데이터를 도서관이 소장, 관리하고, 민간 및 학계와의 지속적인 협력을 기대하고 있다(송민, 2014).

유럽연합(European Union)의 디지털도서관 프로젝트인 유로피아나(Europeana)는 유럽연합 25개 회원국이 보유한 방대한 문화유산을 디지털화하여 서비스하는 프로젝트로 대표적인 도서관 빅데이터 서비스라 할 수 있다. 모바일과 사용자 로그분석을 통해 이용자의 이용패턴을 분석하여 요일별, 시간대별, 체류시간, 페

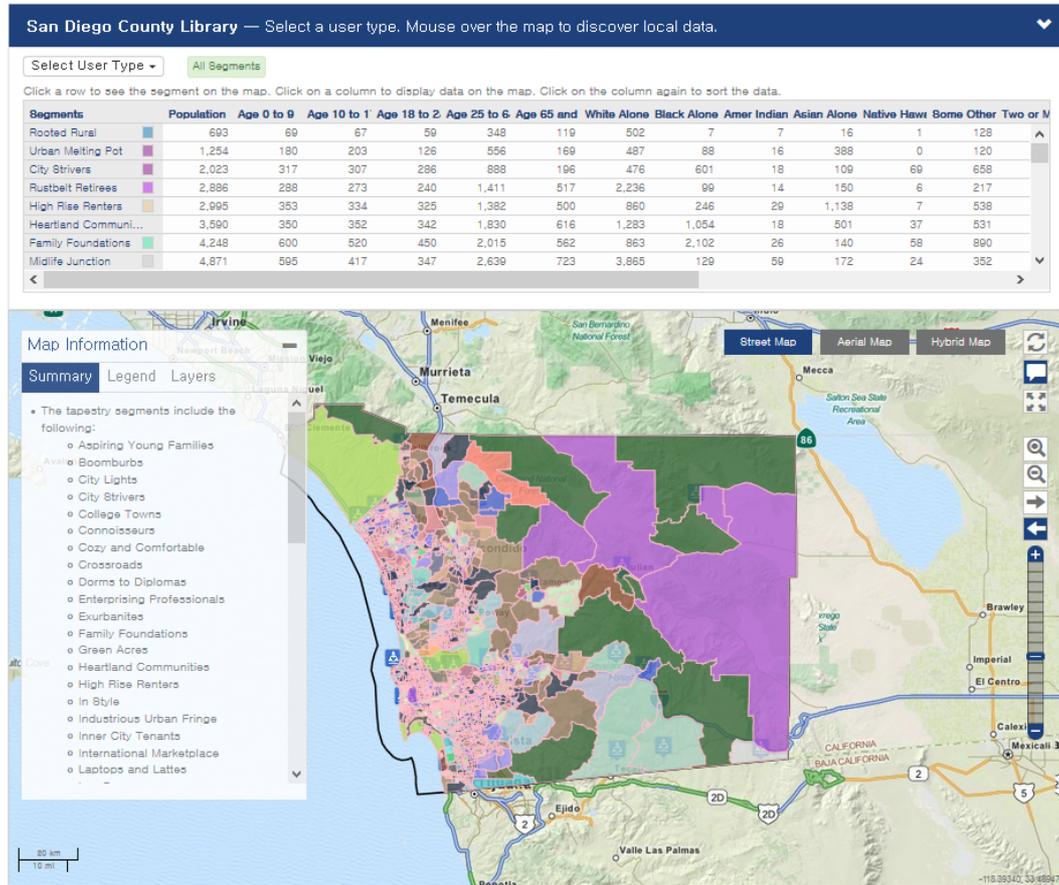
이지뷰, 사용기기를 파악하고, 많이 사용하는 콘텐츠/서비스, 주로 이용하는 시간, OS, 이용자의 국가, 방문시 쿼리, 이용 콘텐츠의 비교, 유입경로, 방문이유, 이용 빈도 등을 파악하여 이용자 서비스 및 콘텐츠, 인프라 강화에 반영하고 있다.

도서관 분야에 지리정보시스템을 적용하여 도서관 경영에 활용하고 있다. 미국의 Community Connect는 도서관 봉사지역의 인구데이터, 학교관련 데이터와 같은 공공데이터 등을 특정 도서관의 장서, 대출, 시설 및 설비관련 데이터 등과 결합하여 도서관 서비스에 활용하고 있다(〈그림 2〉 참조). GeoLib 프로젝트는 미국 내의 공공도서관의 지리적 데이터베이스를 구축하여, 이를 통해 각 개별 공공도서관시스템과 지역을 분석하는 토대로 활용하는 연구로 특정 지역 공공도서관을 중심으로 사방 2마일 내에 있는 지역의 인구 통계적 특성을 보여주는 기능을 제공하고 있다(이성신, 2011).

네덜란드 암스테르담 공공도서관에서는 자료 검색에 정보시각화 기법을 적용하여 관련된 자료를 이용자가 쉽게 인지할 수 있도록 지원하고 있다. 이러한 정보시각화를 통해 이용자들은 정보를 직관적으로 이해할 수 있으며 정보에 친근하게 접근할 수 있게 되고, 데이터의 양이 많아서 의미를 파악하기 어려운 것을 그림으로 표시하여 보여주기 때문에 경향을 쉽게 알 수 있도록 도와줄 수 있다.

국내의 경우, 문화체육관광부는 도서관 건립 계획을 위한 구체적이고 명확한 근거를 기반으로 한 건립계획에 대한 필요성에 의해 '도서관 건립계획 지원시스템'을 개발하였다. 도서관 건립 계획에 따른 예산안을 도출하여 도서관 건립

Connect With Your Community



<그림 2> CommunityConnect에서 시각화 된 서비스 지역의 특성

을 위한 다양한 건립 시뮬레이션을 실행해 봄으로써 이 결과를 상호 비교하여 최적의 도서관 건립계획 수립을 지원하는 기능을 수행한다(문화체육관광부, 2013).

KAIST도서관은 웹로그분석시스템을 구축하여 2013년 1월부터 데이터를 수집하여 도서관 빅데이터를 분석하여 서비스를 개선하였다. 대표적으로 유입 URL 분석으로 포털사이트에서 유입되는 대부분의 이용자들이 '카이스트 도서관'을 키워드로 검색하여 홈페이지에 접근하

는 것을 확인하고, 기존에 사용하던 '한국과학기술원 과학도서관'이라는 명칭 대신 포털사이트에 노출되는 도서관 명칭을 '카이스트 도서관'으로 변경하였다. 또한, 시간대별, 요일별 방문자수 분석을 통해 토요일, 일요일 새벽시간대에 이용자의 검색건수가 가장 적다는 것을 확인하고, 전자도서관 서비스를 중단해야 할 상황이 발생할 경우, 이용량이 적은 토요일, 일요일 새벽 시간대에 서비스를 중단하여 이용자의 불편을 최소화하였다(김서, 2013).

이와 같은 도서관의 빅데이터 서비스의 활용 사례를 서비스 개발 방향의 관점에서 보면 서비스의 수혜 대상이 최종 이용자에 국한 된 것이 아니라 도서관의 정책 결정자, 사서들로 도서관의 경영 및 사서들의 실무와 관련된 빅데이터 서비스가 요구되고 있음을 알 수 있다. 또한, 개발된 서비스들 중 많은 사례가 지리정보 등과 같이 도서관 외부의 공공 데이터와 결합 것을 볼 수 있어 도서관의 빅데이터 서비스는 도서관과 관련된 데이터뿐만 아니라 외부의 공공데이터와 결합하여 개발하는 방향으로 고려할 수 있겠다.

## 2.4 도서관 빅데이터의 가치

### 2.4.1 빅데이터 가치

빅데이터는 그 자체가 중요한 것이 아니라 분석의 결과로 주어지는 가치 부여가 더욱 중요한 문제이다(이정미, 2013). 때문에 빅데이터의 활용은 빅데이터 가치에 대한 인식을 기반으로 한다. 연구 분야와 관계없이 빅데이터의 가치를 보는 관점은 기본적으로 기존의 시스템에서는 한계가 있었던 사안에 대해 좀 더 유용한 정보의 제공 여부, 더 높은 정보 적합성 여부, 더 도움이 되는 의사결정 지원의 여부, 문제 해결의 신속성 향상 여부에서 찾을 수 있다(Villars, Olofson, & Eastwood, 2011).

빅데이터의 가치를 개별 산업 영역에서 구체적으로 보면 빅데이터의 가치는 비용절감과 세수증가(효율성 증대), 생산 혁신, 사업기회의 창출 세 가지 영역에서 가능하다(Centre for Economics & Business Research, 2012). 특히, 빅데이터의 효율성은 6가지의 체제를 통한 기업의 수입과

비용에 대한 영향력을 의미한다.

이 효율성의 6가지 구조는 소비자지성, 공급채널경영, 성과, 품질, 위험관리, 부패방지로 구분할 수 있는데, 빅데이터로 인한 효율성은 각 체제가 영향을 미치는 산업 영역에 따라 다르게 나타난다. 빅데이터로 가능한 혁신은 R&D의 효율적인 이용을 통한 신제품 개발하는 것으로 데이터 중심의 R&D, 장기간에 걸친 사업 효과로 나타난다. 창조는 중소기업에 새로운 시장과 기술의 개발을 위한 장벽 제거에 영향을 줌으로써 고정비용을 감소하고 새로운 시장 개척에 중소기업의 사업 기회를 창출하는 것을 의미한다. 또한, 빅데이터는 새로운 사업기회에 따른 고용을 창출할 수 있는데, 대표적으로 소프트웨어 프로그래머, 데이터 분석가와 같은 새로운 형태의 일자리를 창출할 것으로 보고 있다.

빅데이터의 효과를 분석한 McKinsey Global Institute(2011)는 빅데이터의 가치를 투명성 창출, 요구의 발견, 다양성 도출, 성과 향상, 커스터마이징된 경영을 위한 고객 세분화, 자동화된 알고리즘으로 의사결정 지원과 의사결정의 대체, 새로운 사업 모델, 상품 서비스, 혁신으로 보았다. 생산성이 수량화 된 산출물을 가지고 있지 않고 시간에 따라 변화하는 가치를 반영해야 하는 공공분야에서 빅데이터의 가치를 유럽연합의 공공정보 활용의 관점에서 볼 수 있다. 유럽연합은 공공분야 전반을 통해 본 빅데이터의 정량적인 금전가치와 비금전적 가치를 <표 3>과 같이 구분하여 제시하고 있다.

한국정보화진흥원(2013)은 해외의 다양한 빅데이터 활용 사례 분석을 통해 공공부문 빅데이터 활용의 기대 효과를 예산절감, 정책의 제 선정 근거 확보, 정책 수행 최적화, 대국민

〈표 3〉 유럽연합의 공공분야 빅데이터의 가치

가치의 속성	내 용
정량적 가치	<ul style="list-style-type: none"> <li>• 운영 효율성에서 오는 비용 절감</li> <li>• 행정운영에 있어 허위나 오류의 방지</li> <li>• 세수의 증가</li> </ul>
비정량적 가치	<ul style="list-style-type: none"> <li>• 프로그램에 대한 자원분배의 개선</li> <li>• 높은 품질의 서비스</li> <li>• 증가된 회계 책임성</li> <li>• 더 똑똑한 시민</li> <li>• 정부에 대한 신뢰 향상</li> </ul>

서비스 만족도 제고, 총체적 사회 비용 절감, 공공 기관과 국가 투명성 제고, 위기 대응 능력 향상 등으로 꼽았다. 강만모, 김상락, 박상무(2012)는 빅데이터의 기회 요인을 미래 사회가 갖는 불확실성에 대한 통찰력, 위협요인에 대한 대응력, 스마트 사회에 대한 경쟁력, 융합의 시대에 창조력으로 보았다. 이러한 빅데이터의 기회 요인은 우리 사회의 각 분야에서 찾을 수 있다.

2.4.2 도서관 빅데이터 서비스의 가치  
도서관 빅데이터의 가치는 이용자 정보추구

행동에 기초한 각종 데이터를 신속하고 손쉽게 수집, 분석하여 다양한 정보서비스 개발 및 정책에 반영하여 양질의 정보서비스를 적시적소에 신속하게 제공함으로써 얻을 수 있다. 특히 공공도서관은 공공 분야의 빅데이터 가운데 하나로써 다양한 빅데이터의 효과를 추구할 수 있는데 Centre for Economics and Business Research(2012)가 제시한 빅데이터 서비스의 가치, 한국정보화진흥원(2012)에서 각 분야 빅데이터 서비스를 통해 얻을 수 있는 효과 등을 토대로 하여 공공도서관이 찾을 수 있는 빅데이터의 가치를 도출하면 다음과 같다.

〈표 4〉 도서관 빅데이터 서비스의 가치

가치의 범주	가치의 내용	비고
효율성	<ul style="list-style-type: none"> <li>• 운영 효율성에서 오는 비용 절감</li> <li>• 프로그램에 대한 자원분배의 정책 의제 선정 근거 확보</li> <li>• 정책 수행 최적화</li> </ul>	업무 시간 절감, 중복 예산 절감
혁신	<ul style="list-style-type: none"> <li>• 자동화된 알고리즘으로 의사결정 지원 및 대체</li> <li>• 커스터마이징된 경영을 위한 고객 세분화</li> <li>• 요구의 발견 및 다양성 도출</li> <li>• 성과 향상 및 높은 품질 서비스</li> </ul>	자동화 된 봉사 대상 지역 특성 데이터 추출
창조	<ul style="list-style-type: none"> <li>• 요구의 발견 및 다양성 도출</li> <li>• 신규 서비스 모형 창출</li> </ul>	이용자 대출 기록 분석
향상	<ul style="list-style-type: none"> <li>• 서비스에 대한 신뢰도 향상</li> <li>• 서비스의 질 향상과 만족도 향상</li> </ul>	총체적 서비스 향상 및 만족 증가

첫째, 도서관 업무의 효율성 향상이다. 효율성은 빅데이터 서비스로 인해 얻을 수 있는 직접적인 효과로 시간과 비용의 절감을 의미한다. 이러한 효율성은 Mckinsey Global Institute (2011)의 빅데이터 서비스 가치에서 보면 자동화된 알고리즘으로 의사결정을 지원하거나 대체하거나 프로그램에 대한 자원 분배의 개선을 통해 얻을 수 있는 운영 효율에 따른 비용의 절감이다. 이미 개발되어 서비스 되고 있는 Community Connect나 '도서관 건립계획 지원시스템'은 자동화된 알고리즘으로 도서관의 다양한 정책결정을 지원하며, 본 연구에서 개발한 '수서 추천 기능' 또한 사서의 수서 업무에 시간 및 비용 절감의 효율성이 큰 가치로 간주될 수 있다.

둘째, 도서관 서비스의 혁신이다. 이용자의 집적된 정보요구와 정보이용 데이터는 다양하고 새로운 요구를 말해주는 정보이다. Community Connect나 GeoLib와 같은 도서관 이용자와 공간 및 지리 데이터를 결합한 빅데이터 서비스는 지역 특성과 도서관 이용과의 관계를 보여줌으로써 커스터마이징된 경영을 위해 고객을 세분화할 수 있다. 이러한 도서관 이용자에 대한 정확한 분석은 적극적인 도서관 마케팅을 통해 다양한 지역사회의 요구를 발견, 도출하여 새로운 서비스 모델 개발을 가능하게 한다.

셋째, 창조적인 도서관 서비스의 발굴이다. 도서관 빅데이터를 분석함으로써 그동안 드러나지 않았던 사서 및 이용자들의 다양한 요구를 발견하고, 이를 통해 새로운 서비스 모형을 창출할 수 있다.

넷째, 향상된 서비스 제공이다. 도서관 빅데이터의 가치는 도서관 서비스 품질 향상, 이용자 만족과 편익의 증가를 가능하게 한다. 빅데이터 서비스의 궁극적인 목적은 서비스 수혜자의 최대 만족과 혜택을 높이는 것이다.

### 3. 도서관 빅데이터 서비스 개발

#### 3.1 도서관 빅데이터 서비스 개발 환경

빅데이터 서비스 개발은 공공도서관의 이용자 및 서비스 환경 분석을 통해 핵심 이용자 및 수요자 정의, 수요자에 따른 요구사항 분석, 요구되는 빅데이터 서비스 모형 도출의 과정으로 이루어졌다(〈그림 3〉 참조).

국내외 도서관 분야의 빅데이터 활용사례를 분석한 시사점을 기반으로 공공도서관 빅데이터 서비스의 실질적인 수요자는 도서관의 정책 실무자, 사서, 이용자로 정의하였다. 도서관 정



〈그림 3〉 도서관 빅데이터 서비스 모형 도출 과정

책 실무자는 도서관의 이용 편의를 위한 시스템 또는 비(非)시스템 분야의 정책 수립 의사결정을 담당하는 역할을 수행하는 대상으로 주로 도서관 관장이 해당한다. 둘째, 도서관 사서는 도서관 내부 운영 및 이용자의 편의를 위해 도서관 업무를 수행하는 담당자로, 도서관 사서의 업무는 수서업무, 정리업무, 정보봉사업무, 프로그램업무 등으로 구성된다. 셋째, 도서관 이용자는 도서의 검색, 대출·반납 등 도서관에서 이용자 편의를 위해 제공하는 서비스 수혜자로, 도서관을 이용하는 어린이, 청소년, 주부, 노약자 등 다양한 계층들로 구성된다.

도서관 빅데이터 서비스를 개발하기 위한 서비스 환경 및 요구사항 분석을 위해 도서관 사서 및 관장으로 구성된 실무 전문가와 도서관 분야의 연구를 수행하고 있는 문헌정보학 교수로 구성된 연구자를 대상으로 전문가 심층인터뷰를 통해 도서관의 현황 및 문제점 분석을 수행하였으며, 서비스 수요자별 심층 인터뷰의 주요 내용은 <표 5>와 같다.

심층인터뷰를 통해 도출된 도서관 주요 업무별 현황 및 문제점은 수서, 정리, 열람봉사, 프로그램이라는 공공도서관의 주요 업무별로 파악하였다(<표 6> 참조).

파악된 업무 분야별 현황과 문제점을 통해 빅데이터 서비스의 개발 방향을 수립할 수 있었다. 수서 업무에서는 표준화된 선정 지원도구를 제공함으로써 사서 개인의 역량에 따른 선정업무의 지원을 가능하게 하는 것, 정리 업무에서는 일괄적으로 제공되는 도서 분류기호와 서가배치를 특성화된 장서배치가 가능한 지원도구를 제공하는 것으로 서비스 개발 방향을 설정하였다.

공공도서관의 현황, 문제점, 개발 방향성을 기반으로 빅데이터 서비스 모형을 도출하기 위해 전문가 자문회의와 FGI를 통해 도서관 빅데이터 서비스 모형들을 도출하였다. FGI를 수행한 전문가 그룹은 도서관 정책과 사서의 2개 그룹으로 구분하여 진행하였다. 정책 그룹은 도서관 관장 및 문헌정보학과 교수 2명, 사서 그룹은 공공도서관뿐 아니라 대학도서관, 구립도서관의 사서들 4명으로 구성하여 되도록 다양한 도서관 환경의 의견 수렴을 위해 그룹 내 구성원을 다양하게 포함시켰다. 전문가 자문회의 및 FGI를 통해 우선적으로, 도서관 빅데이터 서비스 모형을 16개 선정하였고 2차로 9개의 모형으로 요약, 최종 4개의 3차 서비스 Pool을 선정하여 도서관 빅데이터 서비스 모형을 도출하였다.

<표 5> 서비스 수요자 심층 인터뷰 내용

구분	대상	인터뷰 주요 내용
실무 전문가	도서관 사서	도서관 사서의 주요 업무, 도서관 운영 애로사항 및 개선 방안, 도서관 사서 전문성 강화를 위한 개선 사항
	도서관 관장	공공도서관 역할, 도서관 서비스 우선순위, 신규 서비스 도입 의사결정에 미치는 요인 등
연구자	문헌정보학과 교수	도서관 의사결정의 애로 사항, 도서관 의사결정 지원을 위한 빅데이터의 활용, 도서관 사서 업무 효율화 및 전문성 향상을 위한 빅데이터 활용, 도서관 빅데이터 서비스의 실질적 도입 및 개발을 위한 기술적, 정책적 요구사항
공통	실무 전문가, 연구자	도서관 분야 빅데이터 도입에 대한 의견 및 서비스 아이디어

〈표 6〉 빅데이터 서비스 모형 도출을 위한 도서관 현황

구분	현황	문제점	개발 방향
수서	<ul style="list-style-type: none"> <li>• 도서관 구입 시 주로 참고하는 자료는 신문서평 정보, 온라인 서점 메인, 베스트셀러 등으로 도서 소개 및 서평정보가 나와 있는 도구를 활용</li> </ul>	<ul style="list-style-type: none"> <li>• 사서가 선정하는 도서, 주제별 균형을 고려한 도서 선정 등은 사서의 전문성을 요구하는 사항</li> <li>• 담당 사서가 보유한 도서 선정기법에 따라 상이한 차이를 보이고 있어 전문성이 부족한 사서들은 도서 선정에 어려움을 겪고 있음.</li> </ul>	<ul style="list-style-type: none"> <li>• 도서 선정 지원을 위한 표준화된 도구 필요</li> </ul>
정리	<ul style="list-style-type: none"> <li>• 국립도서관에서 제공하는 참조도서 분류(Cataloging)을 기본으로 자관에 맞게 재배치, 재정렬하는 업무를 수행</li> </ul>	<ul style="list-style-type: none"> <li>• 대부분의 도서관에서 MARC 업무를 도서 납품업체에 위탁하여 운영함에 따라 해당 도서관 특성에 맞는 적합한 정리업무가 어려운 상황</li> <li>• 납품업체가 제작해 오는 MARC는 국립중앙도서관의 표준 편목을 따르기 때문에 활용성이 결여됨.</li> </ul>	<ul style="list-style-type: none"> <li>• 도서관 특성에 적합한 서가 배치 지원 도구 필요</li> </ul>
열람봉사	<ul style="list-style-type: none"> <li>• 도서관 이용자의 다양한 요구를 해결해야 하는 업무이기 때문에, 담당자마다 각자의 노하우 및 도구 활용 등의 전문성의 큰 차이를 보이고 있음.</li> </ul>	<ul style="list-style-type: none"> <li>• 현재 기록되어 있지 않기 때문에 열람봉사 업무를 담당하는 사서들이 활용하는 도구에 대한 정보 획득 어려움.</li> </ul>	<ul style="list-style-type: none"> <li>• 전국 공공도서관의 열람봉사업무 데이터의 수집 및 저장이 선행된 후에 빅데이터 서비스 개발이 가능할 것으로 보임.</li> <li>• 열람봉사 노하우 및 활용 도구의 공유 필요</li> </ul>
프로그램	<ul style="list-style-type: none"> <li>• 도서관에서 제공하고 있는 교육·문화프로그램이 최신 독서트렌드를 중심으로 구성되어 프로그램 운영 및 강사 섭외가 이루어지고 있음.</li> </ul>	<ul style="list-style-type: none"> <li>• 프로그램 업무를 담당하는 사서들의 기획력, 섭외능력 수준에 따라 큰 차이성을 보이고 있음. 타 도서관의 프로그램 콘텐츠, 강사정보 등이 공유되고 있지 못함.</li> <li>• 강사의 중복으로 인한 강사섭외의 어려움 발생</li> </ul>	<ul style="list-style-type: none"> <li>• 차별화된 문화프로그램 개발 지원 필요</li> <li>• 강사 개발, 섭외 지원 필요</li> </ul>

마지막으로, 도서관 빅데이터 서비스 모형에 대한 수혜자의 수요를 파악하기 위해 FGI와 설문조사를 동시에 수행하여, 도서관 빅데이터 서비스 모형의 개발 우선순위를 파악하고, 우선 개발 서비스를 최종 선정하였다. 도서관 빅데이터 서비스 모형에 대한 설문조사는 사서 및 연구자 200명, 공공도서관 이용자 300명을 대상으로 실시하였다.

### 3.2 빅데이터 서비스 모형 도출

1차적으로 구성된 16개 서비스 모형은 〈표

7〉과 같다. 도서관 정책 결정자를 위한 서비스가 3건, 실무 사서를 위한 서비스가 8건, 이용자 서비스가 5건으로 도서관 사서를 위한 빅데이터 서비스가 가장 많이 제안되었다.

도서관 정책 결정자를 통해 도출된 서비스들은 도서관 건립과 같은 중요한 정책 결정에 근거가 되는 것이나 개별 도서관의 경영과 관련된 결정을 지원하는 것으로 나타났다. 사서들은 실무에 직접적으로 도움이 되는 서비스를 이용자 측면에서는 도서 이용을 촉진할 수 있는 서비스들이 제안되었다.

〈표 7〉 도서관 빅데이터 서비스 모형(1차 16개 서비스)

순번	서비스명	내용	비고
1	실시간 도서관 통계정보 제공 서비스	전국 도서관의 일일/주간 회원가입자수, 방문자수, 대출권수 등의 정책 실무자가 관심 있어 하는 통계정보를 실시간으로 확인·체크할 수 있는 웹기반 서비스 또는 모바일 애플리케이션 서비스	도서관 정책 결정자
2	도서관 입지선정지원 서비스	도서관 건립 예정 지역의 도서관 이용자 거주지 정보, 도서관 이용 동선 데이터(빅데이터)를 분석하여 신규 도서관 건립을 위한 도서관 입지선정의사결정을 지원하는 서비스	
3	VOC 분석을 통한 이용자 서비스 개선	전국 도서관 홈페이지에 게재되는 VOC(Voice of Customer)정보를 수집·분석하여 도서관 이용자의 서비스 개선을 지원하는 서비스	
4	이용패턴 분석을 통한 차별화 전략 도출	특정 도서관에 소장되어 있는 도서, 소장 자료, 기타 간행물 등에 관한 이용자들의 행태분석을 통한 이용패턴을 도출하여 해당 도서관의 특정 및 차별화요소 도출에 적용하는 서비스	
5	도서선정지원_목록분석	전국 도서관에서 선정·구입한 목록(빅데이터)을 분석하여 특정 도서관의 도서 보유현황과 비교하여 도서선정을 지원하는 서비스	도서관 사서
6	도서선정지원_로그분석	전국 도서관의 이용자 검색 로그(빅데이터)를 분석하여 특정 도서관의 도서 보유현황과 비교·분석하여 도서선정을 지원하는 서비스	
7	도서소개 및 추천	전국 도서관에서 수집, 생성되고 있는 도서 소개 및 서평정보(빅데이터)를 수집·분석하여 수서업무를 담당하는 사서의 도서선정 추천 서비스	
8	대출 분석 및 도서 추천	전국 도서관의 도서 구입 및 배치 현황(빅데이터)을 분석하여 이용자들이 원하는 도서를 구입, 배치하데 도움을 주는 서비스	
9	사서업무공유 공개 포털	사서들의 업무 효율화와 사서 간 전문성의 차이를 줄이기 위해 전국 사서들이 해당 업무를 효율적으로 수행하는데 활용하고 있는 각종 도구(도서 선정기법, 자료의 정리 노하우, 자료의 빠른 검색기법 등)를 공유할 수 있는 빅데이터의 공개 포털 서비스	
10	평생교육, 문화행사 추천 서비스	전국 도서관의 평생학습·문화행사 프로그램 콘텐츠 및 강사, 후기정보를 분석하여 특정 도서관의 지역적 특성과 부합하는 적절한 도서관 프로그램을 추천해주는 서비스	
11	도서배달 최적경로지원 서비스	도서관 상호반납 서비스를 운영하고 있는 도서관을 대상으로 도서관 대출·반납 교통량을 파악하여 도서관 배달업무의 최적경로(Critical Path)를 제공해주는 서비스	
12	도서검색 추천 서비스_검색로그	전국 도서관의 이용자 검색 로그(빅데이터)를 분석하여 '이 책과 함께 검색한 도서' 목록을 제공하는 도서검색추천서비스	도서관 이용자
13	대출도서 추천	이용자 대출데이터를 분석하여 신규대출도서를 추천해주는 서비스	
14	연관검색어 추천 서비스	전국 도서관의 이용자 검색로그를 분석하여 검색어 간 연관성을 분석하여 연관검색어를 추천해주는 서비스	
15	지도를 활용한 도서 위치확인 서비스	GIS맵과 연동하여 상호대차를 신청한 도서의 실시간 배송 현황 및 위치를 확인할 수 있는 이용자 편의 서비스	
16	날씨정보에 따른 도서추천 서비스	날씨정보에 따른 도서관 이용현황 분석결과를 바탕으로 날씨에 따라 도서관 이용자에게 맞춤형 도서를 추천해주는 서비스	

### 3.3 도서관 빅데이터 서비스 모형 개발

#### 3.3.1 서비스 모형 선정 과정

개발할 서비스 선정을 위해 도출된 16개의

도서관 빅데이터 서비스 모형을 대상으로 2차 전문가 자문 및 실무자 검토 회의를 진행하였다. 2차 서비스 모형 도출을 위해 16개의 1차 서비스 모형을 코드화하고, 전문가 자문 및 실

무자 검토를 통해 추가, 삭제 등을 판단하여 도출하였다. 2차 서비스 모형은 1차 서비스 모형을 기반으로 유사한 특성의 서비스를 통합, 추가 하여 9개의 서비스로 구성하였고 정책 그룹과 도서관 그룹으로 구성된 전문가 FGI를 통해 <그림 4>와 같이 최종 4개의 도서관 빅데이터 서비스 모형을 도출하였다.

최종적으로 도출된 네 가지 서비스 모형은 다음과 같다.

첫째, 사서 툴킷 서비스는 전국도서관의 도서 선정·구입 목록과 도서소개 및 서평정보를 수집, 분석하여 도서 선정구매 업무를 지원하는 '수서업무지원' 기능과 전국 도서관의 서가배치 데이터와 이용자 검색 데이터를 수집, 분석하여 서가배치 업무를 지원하는 '정리업무지원' 기능을 포함하는 서비스이다. 둘째, 스마트 대시보드 서비스는 도서관 신규 서비스 개발 및 이용자 편의를 위한 의사결정을 지원하는 '의사결정지원' 기능과 상위 도서관 정보 및 타 도서관

비교분석 정보를 차트 형식으로 표출하는 '통계정보제공' 기능을 포함하는 서비스이다. 셋째, 디지털 큐레이터 서비스는 전국 도서관 이용자의 대출데이터와 이용자 검색 로그 데이터를 수집, 분석하여 이용자 맞춤형 추천 서비스를 제공하는 '이용자 맞춤형 추천' 기능과 GIS맵과 연동하여 지도 기반 도서관 정보를 제공함에 따라 도서관 이용자의 이용 편의를 제공하는 'GIS융합정보제공' 기능을 포함하는 서비스이다. 마지막으로 도서관 프로그램 프로파일링 서비스는 전국 도서관의 평생교육, 문화행사 프로그램 관련 데이터를 수집, 분석하여 인기강좌, 인기강사를 분석하는 '인기강좌, 인기강사 분석' 기능과 전국 공공도서관의 문화프로그램을 수집하여 문화프로그램 단위로 도서관 문화 프로그램을 비교·분석하는 '도서관별 문화 프로그램 비교' 기능을 포함하는 서비스이다.

우선적으로 개발할 서비스를 서비스 개발의 필요성, 시급성, 개발 가능성 관점에서 다시 검토



<그림 4> 도서관 빅데이터 서비스 모형의 도출

하였다. 도서관 사서 및 문헌정보학과 교수들을 활용한 자문회의뿐만 아니라 도서관 사서와 이용자 대상 설문조사로 구체적인 요구를 조사함으로써 서비스 모형 개발을 위해 선정된 최종 4개의 빅데이터 서비스의 내용을 구체화 하였다. 설문지는 도서관 서비스 만족, 서비스의 관점에서 수서 및 정리 업무의 만족, 도출된 빅데이터 서비스 모형에 대한 수요도 등으로 구성되었다. 2014년 7월 21에서 30일 약 2주에 걸친 설문조사는 사서 200명, 이용자 300명을 대상으로 하였다. 설문 당시 시점에서 도서관의 빅데이터 서비스에 대한 인식이 낮은 점, 실질적인 빅데이터 서비스에 대한 수요 파악을 위한 설문조사라는 점을 고려해 설문조사는 관련 지식에 대한 관심도와 이해가 선행되었다고 판단한 빅데이터 서비스 관련 세미나에 참석한 사서들을 대상으로 하였다. 이용자는 서울시 중심에 위치한 공공도서관 이용자를 대상으로 면대면 수집하였다.

사서를 대상으로 한 설문조사에서는 공공도서관에서 아직까지 도서관 문화프로그램 등에 관한 수요가 높지 않고, 도서관 정책 의사결정권자를 위한 서비스를 개발하는 것보다 도서관

사서 업무의 효율화, 도서관 이용자 맞춤형 서비스에 대한 수요가 더 높은 것으로 나타났다. 설문 결과와 전문가 의견을 바탕으로 우선 개발 서비스를 <표 8>과 같이 4개의 서비스 후보를 선정하였다.

선정된 네 가지 서비스 모형을 <표 7>의 1차 16개 서비스 모형과 연결시키면 다음과 같다. 첫째, 사서툴킷서비스는 '5. 도서선정지원\_목록 분석,' '6. 도서선정지원\_로그분석,' '7. 도서소개 및 추천'에 해당하며, 둘째, 스마트 대시보드 서비스는 '1. 실시간 도서관 통계정보 제공 서비스'를 의미한다. 셋째, 디지털 큐레이션 서비스는 이용자를 위한 '도서관 이용자를 위한 '8. 대출 분석 및 도서 추천'이며, 마지막으로, 도서관 프로그램 프로파일링 서비스는 도서관 문화 프로그램 업무를 지원하는 '10. 평생교육, 문화행사 추천서비스'에 해당한다.

전체적인 이용자의 빅데이터 서비스 수요와 함께 우선 개발 서비스로 선정된 사서 툴킷 서비스와 디지털 큐레이터 서비스를 좀 더 구체적으로 구현하기 위한 목적으로 이용자를 대상으로 한 설문조사를 수행하였다. 설문은 이용자의 도서관 이용, 도서관 만족, 빅데이터 서비

<표 8> 도서관 빅데이터 서비스 우선순위 결정(사서)

순번	서비스명	내용	필요성	시급성	개발 가능성	비고
1	사서 툴킷 서비스	도서관 내부 운영을 위한 사서업무를 지원하는 서비스	●	●	●	우선 개발
2	스마트 대시보드 서비스	도서관 정책 의사결정을 지원하는 의사결정 지원 서비스	●	⦿	⦿	중장기 개발
3	디지털 큐레이터 서비스	도서관 이용자의 도서 열람, 대출 편의를 위한 맞춤형 추천 서비스	●	●	●	우선 개발
4	도서관 프로그램 프로파일링 서비스	도서관 프로그램 업무의 기획, 운영업무를 지원하는 서비스	⦿	○	⦿	중장기 개발

●: High, ⦿: Medium, ○: Low

스에 대한 이용자들의 의견에 대한 항목으로 구성하였다.

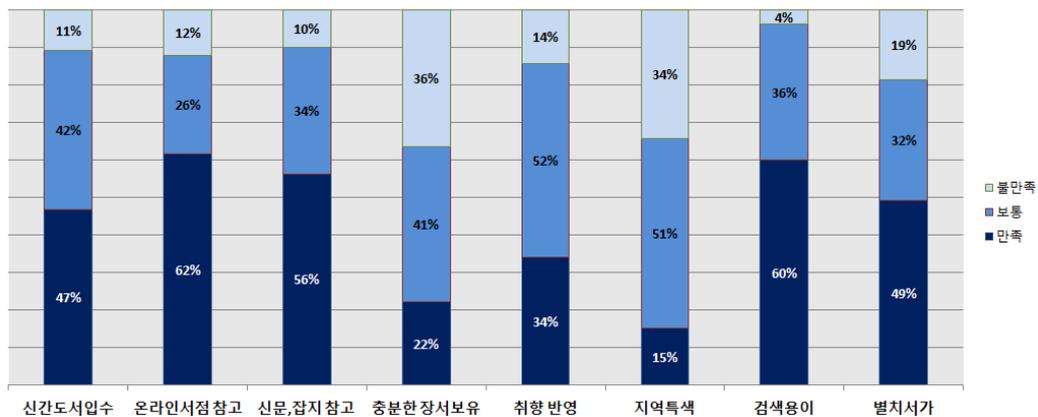
설문 결과를 간략히 정리하면, 이용자들이 도서관을 이용하는데 영향을 주는 요인은 도서관의 위치와 장서이며, 도서 대출을 위해 참고하는 것은 신문, 잡지 등 대중 매체나 전문가들의 추천도서 등으로 나타났다. 또한 이용자들은 도서의 검색 및 찾기, 사서 응대에는 만족도가 높으나 지역특성에 대한 차별성, 최신의 정보기술을 반영한 서비스에 대해서는 낮은 만족도를 보였다. 또한 이용자들은 의 잠재적 요구를 수서 및 정리 업무, 열람 및 대출과 관련한 만족도를 통해 종합하면 충분한 장서 확충이 가장 시급한 것으로 나타났다.

장서 확충에 대한 요구를 <표 8>의 사서툴킷 서비스와 디지털 큐레이션 서비스의 개발의 측면에서 조금 더 구체적으로 살펴보면 <그림 5>, <그림 6>과 같다.

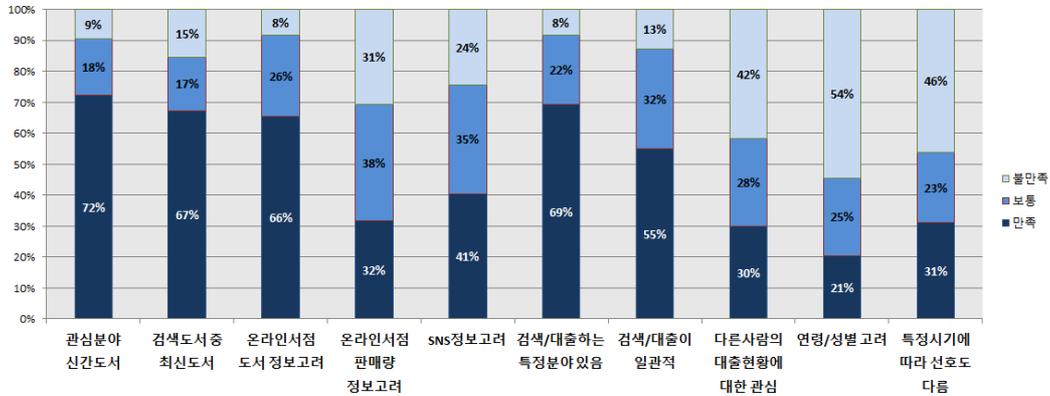
먼저 사서들의 도서 구매 및 서가 배치와 관련한 조사에서 불만족이 가장 높은 항목, 즉 개선 및 수요가 높은 상위 3개를 보면 충분한 장

서의 보유(불만족 36%), 지역 특색을 나타내는 장서(불만족 34%) 별치서가(불만족 19%)의 순으로 나타났다. 이와 같은 결과는 사서들은 빅데이터 서비스를 통해 이용자들이 읽고자 하는 도서를 언제든지 제공할 수 있는 충분한 장서의 보유나 연계서비스를 원하고 있으며, 지역의 특색을 반영한 장서의 보유와 별치서가를 통한 손쉬운 정보 서비스에 대한 요구가 많은 것으로 나타났다(<그림 5> 참조).

도서관 이용자 측면에서 불만족이 가장 높은 상위 3개의 항목은 연령과 성별에 따른 이용(불만족 54%), 특정 시기에 따라 선호도가 다름(불만족 46%), 다른 사람의 대출에 대한 관심(불만족 42%)이다(<그림 6> 참조). 즉, 이용자는 대출 시 별치 서가의 책을 고려하기 보다는 자신의 관심사에 맞는 특정 분야의 도서, 자신과 유사한 관심을 갖는 다른 이용자의 대출 도서에 관심이 많은 것으로 나타나 앞에서 선정한 디지털 큐레이션 서비스가 개인에게 특화된 맞춤형 도서 추천서비스를 이용자 서비스의 주요 모형으로 도출한 것과 일치한다.



<그림 5> 선정된 서비스를 기반으로 한 빅데이터 서비스 도출을 위한 요구(사서)



〈그림 6〉 선정된 서비스를 기반으로 한 빅데이터 서비스 도출을 위한 요구(이용자)

이와 같은 과정을 거쳐 개발을 위해 최종 선정된 도서관 빅데이터 서비스는 도서관 정책결정자 및 사서의 업무에 도움을 줄 수 있는 사서 툴킷 서비스, 스마트 대시보드 서비스와 도서관을 이용하는 이용자들을 위한 디지털 큐레이터 서비스 세 가지이다.

### 3.3.2 최종 개발 서비스 모형: 사서 의사결정 지원 서비스와 이용자 도서 추천 서비스

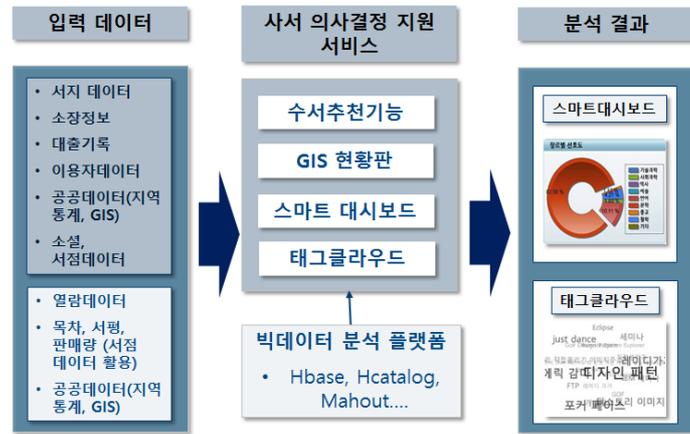
사서 의사결정 지원 서비스는 사서 툴킷 서비스와 스마트 대시보드 서비스를 통합하여 지칭한 서비스명이다. 사서 의사결정 지원 서비스는 자관의 미보유 도서 중 타관 대출빈도가 높은 도서 및 보유 장서를 비교, 분석하여 수서 목록으로 추천하는 수서추천 기능, 도서관 핵심 이용자 및 도서 및 시간대 요일별 대출현황

을 수집 분석하여 통합적으로 대시보드 방식으로 표출해주는 스마트 대시보드 기능, 지역도서관의 규모, 대출 현황 및 소득, 학력, 다문화 등의 인구 특성을 수집 분석하여 GIS 맵 기반으로 쉽고 편하게 정보를 제공해주는 GIS 현황판 등을 주요 기능으로 한다. 이를 개념화 하면 〈그림 7〉과 같이 나타낼 수 있다.

사서 의사결정 지원 서비스의 수서 추천 서비스는 전국 공공도서관의 장서데이터, 대출기록, 이용자데이터와 온라인 서점의 도서 소개 및 서평정보를 수집, 분석하여 수서업무를 지원하는 업무 지원용 빅데이터 서비스이다. 자관 특성에 적합한 구입도서 선정에 도서를 추천함으로써 수서 업무를 지원하고자 한다. 구입 도서 선정을 위한 추천도서는 이용자별 도서 대출·열람 현황, 자관의 장서 보유 현황과 타

〈표 9〉 최종 개발을 위해 선정된 도서관 빅데이터 서비스

순번	서비스명	내용
1	사서 툴킷 서비스	도서관 내부 운영을 위한 사서업무를 지원하는 서비스
2	스마트 대시보드 서비스	도서관 정책 의사결정을 지원하는 의사결정 지원 서비스
3	디지털 큐레이터 서비스	도서관 이용자의 도서 열람, 대출 편의를 위한 맞춤형 추천 서비스



〈그림 7〉 사서 의사결정 지원 서비스의 개념도

관의 장서 보유 현황 비교 분석, 온라인 서점의 베스트셀러 및 스테디셀러 비교 분석, 각종 기관의 추천 도서의 분석을 통해 원하는 기간 및 조건으로 검색하여 선정 도서 목록에 올릴 수 있도록 한다.

스마트 대시보드는 도서관의 주요한 의사결정을 지원하기 위한 서비스로 도서관 핵심 이용자, 시간대별 요일별 대출 현황, 도서관 이용자 행태 분석과 같은 도서관 이용 데이터를 집

계하는 기능으로 다양한 도서관 서비스 및 정책 수립과 결정에 근거가 되는 기본 데이터를 제공하는 기능을 한다. 이러한 의사결정을 위한 서비스는 지역사회 현황을 시각화 하여 제시하는 GIS 현황판, 태그 클라우드와 같은 기능을 통해서도 제공할 수 있다. 이러한 기능들은 〈표 10〉과 같이 크게 4가지로 수서 추천 기능, 스마트 대시보드, GIS 현황판, 태그 클라우드로 구분하였다.

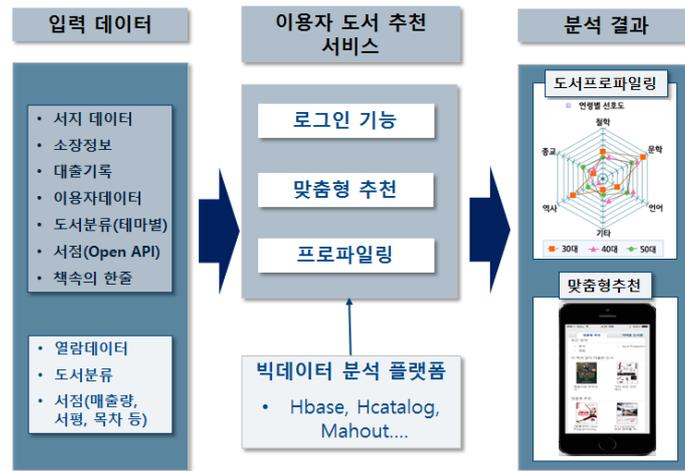
〈표 10〉 사서 의사결정 지원 서비스 주요 기능

구분	정의	예시
수서 추천 기능	• 수서업무 및 정리업무 지원 기능을 통합하여 도서관 운영에 도움을 주는 기능	• 자관 미보유 도서 중 타관 대출빈도가 높은 도서 목록 추천 • 복본 구매도서 추천 • 언론사, 서점의 추천도서 목록 • 서가배치 추천
스마트 대시보드	• 도서관 정책 의사결정자를 위한 도서관 데이터 집계 기능	• 도서관 핵심 이용자/도서 분석 (시간대, 요일별 대출 현황) • 도서관 이용자 행태 분석
GIS 현황판	• 지도 기반으로 도서관 정보를 쉽고 편하게 제공하는 기능	• 지역도서관 현황(규모, 대출 등) • 인구 특성(소득, 학력, 다문화 등)
태그 클라우드	• 빅데이터 분석 결과를 더욱 명확하게 표현할 수 있는 대표적인 빅데이터 가시화 기능	• 최신 도서 트렌드 분석

이용자 도서 추천 서비스 서비스는 도서관 이용자의 대출데이터와 이용자 검색 로그데이터, 잠재적 요구를 수집, 분석하여 이용자 맞춤형 도서 서비스를 추천 제공하는 서비스이다. 즉, 도서관 내부에서 보유하고 있는 이용자 데이터와 도서관 이용자 서비스를 이용하여 생성되는 데이터를 활용하여 도서 검색 및 도서관 이용 편의를 제공하는 맞춤형 도서 추천 빅데이터 서비스이다. 이 서비스는 이용자가 자신의 독서 기록을 저장하면 다른 이용자들의 독서 기록 공공 도서관에서 수집한 지역, 성별, 연

령대별 대출기록 등을 비교, 분석하여 개별 이용자들에게 적합한 도서 프로파일링, 맞춤형 추천 결과를 제공한다(〈그림 8〉 참조).

이용자 도서 추천 서비스 목표는 첫째, 전국 공공도서관에서 보유하고 있는 대출 데이터를 기반으로 성별, 연령, 지역 등을 중심으로 맞춤형 도서를 추천하여 도서관 이용의 편의를 제공하는 것이다. 둘째, 도서관 이용자 데이터 서비스를 통해 수집 및 저장되는 개인별 소장 도서와 도서관 대출데이터와 연계하여 독서 프로파일링 분석으로 균형 있는 도서 대출 및 독서



〈그림 8〉 이용자 도서 추천 서비스의 개념도

〈표 11〉 이용자 도서 추천 서비스 주요 기능

구분	정의	예시
로그인 기능	• 이용자 도서 추천 서비스에 접속하기 위한 로그인 방식	• 자체 로그인 • SNS 연계 로그인
맞춤형 추천	• 검색어 입력 결과로 나타는 맞춤형 추천 기능	• 개인 특성별 신간도서 추천 • 특정 도서별 군집화 목록 추천 • 이용자 독서 성향별 추천
프로파일링	• 연령, 성별, 개인별 독서 취향을 분석해 주는 기능	• 카테고리별 상위 30권 • 개인 독서 프로파일링

〈표 12〉 도서관 빅데이터 서비스를 위한 활용 데이터(예시)

구분	사서 의사결정 지원 서비스	이용자 도서 추천 서비스
정형 데이터	<ul style="list-style-type: none"> <li>• 서지데이터, 소장정보 데이터</li> <li>• 대출기록 데이터, 이용자 데이터</li> <li>• 열람기록 데이터</li> <li>• 온라인 서점의 도서 판매량</li> </ul>	<ul style="list-style-type: none"> <li>• 서지데이터, 소장정보 데이터</li> <li>• 대출기록 데이터, 이용자 데이터</li> <li>• 검색로그 데이터</li> <li>• 열람기록 데이터</li> <li>• 온라인 서점의 도서 판매량</li> </ul>
반정형 데이터	<ul style="list-style-type: none"> <li>• 구글 트렌드 정보</li> <li>• 공공데이터(날씨, 지역통계, GIS 등)</li> </ul>	<ul style="list-style-type: none"> <li>• 구글 트렌드 정보</li> <li>• 공공데이터(날씨, 지역통계, GIS 등)</li> <li>• 도서관 데이터(위치정보)</li> </ul>
비정형 데이터	<ul style="list-style-type: none"> <li>• 도서 목차, 서평</li> <li>• 소셜, 서점 데이터</li> </ul>	<ul style="list-style-type: none"> <li>• 도서 목차, 서평정보</li> <li>• 소셜, 서점 데이터</li> </ul>

환경을 제공하고자 하는 것이다. 이용자 도서 추천 서비스에서 제공하는 주요 기능은 크게 3가지로 로그인 기능, 맞춤형 추천, 프로파일링 등이며 상세 내용은 〈표 11〉과 같다.

위의 빅데이터 서비스 구현을 위해 활용되는 데이터는 〈표 12〉와 같이 정형·반정형·비정형 데이터로 구분한다. 기존에 도서관이 보유하고 있는 서지데이터, 소장정보, 대출기록, 이용자데이터, 공공데이터, 소셜/서점데이터 및 향후 확보해야 할 열람데이터, 목차, 서평, 판매량 등의 서점 데이터와 공공데이터를 입력한다. 빅데이터 분석 플랫폼을 기반으로 제공하는 4가지 주요 기능을 통해 사서들이 이해하기 쉬운 형태의 스마트 대시보드 태그 클라우드 형태의 결과를 표출한다.

#### 4. 결론

본 연구는 최근 많은 이슈가 되고 있는 빅데이터를 도서관 분야에 적용하여 다양한 형태의 도서관 빅데이터의 활용 가치에 대한 이해를 높이고 이에 대한 수요자의 요구 분석을 바탕

으로 공공도서관 빅데이터 서비스 모형을 개발함으로써 변화하는 환경에 맞는 향상된 도서관 서비스를 제공하기 위해 수행되었다.

주요 내용을 보면 첫째, 도서관 빅데이터의 개념과 내용을 고찰하였다. 도서관의 빅데이터는 도서관 정보와 빅데이터의 특성이 결합하여 도서관의 내부 및 외부에서 수집 혹은 생성되는 다양한 도서관 정보로 데이터의 크기, 형태, 생성 속도가 빅데이터의 특성을 모두 포함하여 도서관 분야의 가치를 창출할 수 있는 데이터로 볼 수 있다. 이러한 도서관의 빅데이터는 도서관의 MARC데이터, 대출 정보 등 전통적으로 도서관에서 생성되는 정형데이터, 수집된 정보를 바탕으로 도서관에서 자체적으로 생성되는 반정형데이터로 도서관 통계데이터, 공공정보 등이 포함된다. 또한 게시판, SNS 등을 통해 생성되는 텍스트데이터도 도서관의 비정형데이터로 중요하다.

둘째, 도서관 빅데이터 서비스에 대한 수요 분석을 바탕으로 도서관 빅데이터 서비스 모형을 도출하였다. 도서관 현장의 실무자, 정책 관리자, 연구자, 이용자와 같은 도서관의 핵심 이해관계자 집단의 심층인터뷰, FGI, 설문조사 등

을 통해 1차적으로 16개의 도서관 빅데이터 서비스 모형들을 도출하였다. 이후 최종적으로 서비스 요구에 대한 필요성, 시급성, 개발 가능성 등을 고려하여 두 가지 서비스 모형인 사서의 사결정 지원 서비스와 이용자 도서 추천 서비스를 최종 개발 서비스로 선정하였다. 각 서비스의 상세 개발 방안, 주요 기능 및 시나리오를 제시하였고 서비스 개발에 요구되는 데이터를 기술함으로써 향후 다른 빅데이터 서비스 개발 모형을 제시하였다.

셋째, 도서관 빅데이터의 가치는 향후 서비스 활용으로 예상되는 업무의 효율성, 서비스의 품질 향상, 이용자 만족과 같은 성과를 목표로 할 수 있는 평가의 틀을 제공할 수 있다. 이러한 관점에서 도서관 빅데이터 서비스의 경제적 타당성 분석을 위한 가치평가 항목 및 가치평가 방안을 제시하였다.

본 연구는 그동안 이론, 국내외 동향 및 사례 분석 등에 그쳤던 도서관의 빅데이터를 심층적으로 분석하고 실질적으로 활용할 수 있는 서비스 모형을 도출 및 개발하였다. 본 연구는 공공도서관 분야의 빅데이터 서비스를 활용한 연구 영역을 확장하는 점에서 학술적으로 의미가 있고, 도서관 분야에 빅데이터를 접목한 서비스 개발로 도서관 관련 이해관계자의 다양한 요구를 만족시킬 수 있어 도서관 서비스 및 정책의 발전 및 활성화에 기여할 수 있다고 하겠다. 다만, 본 연구에서는 공공도서관에 적용할 수 있는 표준 모형이고, 의사결정을 지원하기 위한 도구로 사용하는 것이기에 각 도서관은 자관의 상황과 특성에 따라 필요한 데이터를 활용함으로써 올바른 의사결정을 내려야 할 것이다. 또한 공공도서관 및 이용자들로부터 수

집하게 될 다양한 정보 가운데 개인정보를 어떻게 보호할 것인지에 대한 방안도 필요하다. 공공도서관에서 각 이용자들에게 개인정보 수집과 이에 대한 활용, 그리고 제 3자에 대한 동의를 새롭게 받는 것이 현실적으로 불가능하기 때문에 개인정보보호법의 테두리 안에서 개인정보를 수집할 수 있는 방안에 대한 많은 고민과 논의가 필요하다.

본 연구에서는 서비스 도출 및 가치 측정 분야에서 다음과 같은 한계점이 존재하여 향후 연구에서 보완하고자 한다. 첫째, 본 연구를 통해 도출한 도서관 빅데이터 서비스는 도서관 분야의 사서 및 이용자들의 주요 현안을 빅데이터 기술을 활용하여 해결하기 위한 것이다. 즉, 문제 해결형 서비스만을 도출함으로써 빅데이터 및 최신 정보기술을 활용하여 미래 도서관 분야에 적용할 수 있는 미래 지향적 관점의 서비스는 도출하지 못했다. 따라서 빅데이터, 사물인터넷(Internet of Thing: IoT), 클라우드 컴퓨팅 등 최신 정보기술 동향을 조사, 분석하여 도서관 분야에 적합한 미래 지향적 도서관 빅데이터 서비스 모형 수립에 관한 후속 연구가 수행되어야 할 것이다.

둘째, 개발된 빅데이터 서비스의 성과를 본 연구에서 제시한 도서관 빅데이터 가치 항목으로 측정할 필요가 있다. 제시된 가치 항목들은 많은 부분 공공분야에서 측정하기 어려운 정량적 척도로 구성되어 있어 개발되고 사용되는 빅데이터 서비스의 특성과 성과에 따라 구체적인 측정 척도의 적용 방안이 요구된다. 이와 같은 성과에 측정 계획은 서비스 개발과 동시에 이루어짐으로써 개발된 빅데이터 서비스의 목적을 달성해야 할 필요가 있다.

셋째, 도서관의 빅데이터는 그동안 도서관과 관련해 생산된 모든 데이터를 의미하는 것으로 이러한 데이터를 활용한 새로운 서비스의 개발 뿐만 아니라 도서관 데이터의 전반적인 생성, 활용, 성과를 살펴보고, 사회적 관점에서 데이터 활용의 확장에 대한 이론적인 연구가 지속적으로 진행될 필요가 있다. 즉, 데이터 활용의

관점에서 축적된 데이터의 의미, 활용 체계에 대한 전반적인 검토가 필요하다.

마지막으로, 본 연구에서의 서비스 도출 및 선정 방향은 공공도서관을 대상으로 하였기 때문에, 공공도서관에서 빅데이터의 분석과 활용이 성공적으로 완수된다면 이를 기초로 대학도서관이나 학교도서관 등으로 확장되어야 할 것이다.

## 참 고 문 헌

- 강만모, 김상락, 박상무 (2012). 빅 데이터의 분석과 활용. 정보과학회지, 30(6), 25-32.
- 교육과학기술부, 행정안전부, 지식경제부, 방송통신위원회, 국가과학기술위원회 (2012). 스마트 국가 구현을 위한 빅데이터 마스터 플랜.
- 김서 (2013). 빅데이터 도서관 활용사례: KAIST도서관 웹로그분석시스템을 중심으로. KSLA bulletin, 5, 45-52.
- 도서관정보정책위원회 (2014). 제2차 도서관발전종합계획(2014~2018). 세종: 도서관정보정책위원회
- 문화체육관광부 도서관정보정책기획단 [편] (2013). 2012 공공도서관 건립·운영 계획 수립 지원을 위한 웹 시스템 개발 결과 보고서. 서울: 문화체육관광부 도서관정보정책기획단.
- 배동민, 박현수, 오기환 (2013). 빅데이터 동향 및 정책 시사점. 정보통신정책연구, 25(10), 37-74.
- 송민 (2014). 빅데이터 시대, 도서관의 도전과 변화. 도서관 분야 빅데이터 활용 세미나 발표 자료집. 서울: 한국과학기술정보연구원.
- 이성신 (2011). 지리정보시스템을 활용한 공공도서관 마케팅. 정보관리학회지, 28(3), 179-195.  
<http://dx.doi.org/10.3743/KOSIM.2011.28.3.179>
- 이정미 (2013). 빅데이터의 이해와 도서관정보서비스에의 활용. 한국비블리아학회지, 24(4), 53-83.
- 최성곤, 오진태, 장종수 (2012). 미래 지식 서비스를 위한 빅데이터 처리의 기술적 요구사항. 한국정보기술학회지, 10(3), 5-10.
- 한국정보화진흥원 빅데이터연구센터 (2012). 빅데이터로 진화하는 세상: big data 글로벌 선진사례. 서울: 한국정보화진흥원.
- 홍승필, 장현미, 이연우, 김지영, 김동희, 박은충, 손주영 (2012). 빅데이터 기반 개인정보보호 기술수요 분석. 서울: 한국인터넷진흥원.
- Centre for Economics and Business Research (2012). Data equity: Unlocking the value of big data. [cited 2014.9.20]. Retrieved from

- <http://www.sas.com/offices/europe/uk/downloads/data-equity-cebr.pdf>
- Gartner (2012). The importance of big data: A definition. [cited 2012.6.21]. Retrieved from <http://www.gartner.com/doc/2057415/importance-big-data-definition>
- Mckinsey Global Institute (2011). Big data: The next frontier for innovation, competition, and productivity. [cited 2011.3]. Retrieved from [http://www.mckinsey.com/insights/business\\_technology/big\\_data\\_the\\_next\\_frontier\\_for\\_innovation](http://www.mckinsey.com/insights/business_technology/big_data_the_next_frontier_for_innovation)
- IDC (2011). Extracting value from chaos. [cited 2011.6]. Retrieved from <http://www.emc.com/collateral/analyst-reports/idc-extracting-value-from-chaos-ar.pdf>
- Villars, R. L., Olofson, C. W., & Eastwood, M. (2011). Big data: What it is and why you should care. White Paper. IDC. [cited 2015.2.5]. Retrieved from [http://sites.amd.com/us/Documents/IDC\\_AMD\\_Big\\_Data\\_Whitepaper.pdf](http://sites.amd.com/us/Documents/IDC_AMD_Big_Data_Whitepaper.pdf)

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

- Bae, D., Park, H., & Oh, K. (2013). Big data trends and policy implications. *Journal of Information and Communication Policy Research*, 25(10), 37-74.
- Big Data Research Center, National Information Society Agency (2012). *The world evolving into big data: Big data global best practices*. Seoul: National Information Society Agency.
- Choi, Seong-Gon, Oh, Jin-Tae, & Jang, Jong-Soo (2012). Big data skills for future knowledge services. *Korean Institute of Information Technology Magazine*, 10(2), 5-10.
- Hong, S., Jang, H., Lee, Y., Kim, J., Kim, D., Park, E., & Son, J. (2012). *Big data privacy technology-based needs assessment report*. Seoul: Korea Internet & Security Agency.
- Kang, M., Kim, S., & Park, S. (2012). Analysis and utilization of big data. *Communications of the Korea Information Science Society*, 30(6), 25-32.
- Kim, S. (2013). Big data library use cases: KAIST library web log analytics system. *KSLA bulletin*, 5, 45-52.
- Lee, J. (2013). Understanding big data and utilizing its analysis into library and information services. *Journal of the Korean Biblia Society for Library and Information Science*, 24(4), 53-83.
- Lee, S. (2011). Applications of geographic information systems in public library marketing. *Journal of the Korean society for information management*, 28(3), 179-195.

- Library and Information Policy Committee (2014). The second library development master plan (2014~2018). Sejong: Library and Information Policy Committee.
- Library and Information Policy Planning Division, Ministry of Culture, Sports, and Tourism (2013). 2012 Web systems development report for the planning to constructing and operating public libraries. Seoul: Library and Information Policy Planning Division, Ministry of Culture, Sports, and Tourism.
- Ministry of Education, Science and Technology, Ministry of Public Administration and Security, Ministry of Knowledge Economy, the Korea Communications Commission, State Science and Technology Commission (2012). Big data master plan for the implementation of smart country.
- Song, M. (2014). Big data era, challenges and changes in the libraries. Proceedings of Seminar of Big Data Applications in the Library Fields. Seoul: KISTI.