

도서관 빅데이터 플랫폼을 활용한 공공도서관 빅데이터 분석 연구: 대전한밭도서관을 중심으로*

Big Data Analysis for Public Libraries Utilizing Big Data Platform: A Case Study of Daejeon Hanbat Library

온정미 (Jeongmee On)**

박성희 (Sung Hee Park)***

초 록

2016년 1월 1일부터 공공도서관 빅데이터 플랫폼이 서비스되기 시작하여 도서관 빅데이터가 공공도서관 업무 개선에 활용되고 있다. 본 논문은 도서관 빅데이터 플랫폼 활용사례들을 살펴보고 도서관 빅데이터 플랫폼의 활용효과를 높일 수 있는 개선방안을 도출하고자 한다. 이를 위해 먼저, 도서관 빅데이터 플랫폼을 활용한 사례들에서 활용한 빅데이터와 활용유형분석 및 도출된 서비스/시행정책을 살펴본다. 다음으로, 현재 공공도서관에서 사용하는 통합도서관리시스템(ILUS)과 도서관 빅데이터 플랫폼 각각의 자료분석 방식을 비교함으로써 도서관 빅데이터 플랫폼의 한계점과 이점을 살펴본다. 사례분석 결과, 프로그램 기획 및 수행, 장서, 수서, 기타의 유형으로 빅데이터를 활용하였고 서비스/시행정책은 이용자 맞춤형 테마서가 및 독서진흥프로그램 진행, 장서활용도 증대, 특화주제에 기반한 수서 및 대출현황 데이터 공개로 요약되었다. 비교분석결과, ILUS는 자관의 자료실현황분석에 특화되어 있으며, 빅데이터 플랫폼은 다양한 속성(연령, 성별, 지역, 대출시기 등)에 따른 선택적 분석이 가능하여 분석시간단축과 유연한 분석이 가능하다. 마지막으로 사례분석과 비교분석에서 밝혀진 특징 및 한계점을 정리하고 개선방안을 제시한다.

ABSTRACT

Since big data platform services for the public library began January 1, 2016, libraries have used big data to improve their work performance. This paper aims to examine the use cases of library big data and attempts to draw improvement plan to improve the effectiveness of library big data. For this purpose, first, we examine big data used while utilizing the library big data platform, the usage pattern of big data and services/policies drawn by big data analysis. Next, the limitations and advantages of the library big data platform are examined by comparing the data analysis of the integrated library management system (ILUS) currently used in public libraries and data analysis through the library big data platform. As a result of case analysis, big data usage patterns were found program planning and execution, collection, collection, and other types, and services/policies were summarized as customizing bookshelf themes for the book curation and reading promotion program, increasing collection utilization, and building a collection based on special topics, and disclosure of loan status data. As a result of the comparative analysis, ILUS is specialized in statistical analysis of library collection unit, and the big data platform enables selective and flexible analysis according to various attributes (age, gender, region, time of loan, etc.) reducing analysis time. Finally, the limitations revealed in case analysis and comparative analysis are summarized and suggestions for improvement are presented.

키워드: 도서관 빅데이터, 빅데이터 플랫폼, 사례연구, 비교분석

library big data, big data platform, case study, comparative analysis

* 본 연구는 한남대학교 일반대학원 문헌정보학전공 석사학위논문을 축약한 것임.

** 대전한밭도서관 자료정책과장(jmon8527@korea.kr) (제1저자)

*** 한남대학교 문헌정보학과 조교수(sunghee@hnu.kr) (교신저자)

■ 논문접수일자: 2020년 5월 27일 ■ 최초심사일자: 2020년 6월 10일 ■ 게재확정일자: 2020년 8월 25일

■ 정보관리학회지, 37(3), 25-50, 2020. <http://dx.doi.org/10.3743/KOSIM.2020.37.3.025>

© Copyright © 2020 Korean Society for Information Management

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

1. 서론

현재 우리는 정보화 시대를 넘어 다양한 데이터로 넘쳐나는 빅데이터 시대에 살고 있다. 빅데이터란 디지털 환경에서 생성되는 대규모 데이터로, 기존의 기술 방식으로는 저장하기 어려운 큰 규모의 자료를 일컫는다. 과거의 H/W 중심에서, 현재의 S/W 중심을 거쳐 점점 데이터 중심의 환경으로 변화하고, 정보기술(IT)의 지속적인 발전과 스마트폰의 대중화, 사물인터넷(IoT)의 발달, 인공지능의 가시화 등으로 빅데이터에 대한 관심은 더욱 증대하고 있다. 빅데이터는 그 활용성과 가치에 대해 기업은 물론 공공기관, 의료, 방송, 문화, 교육, 과학분야 등 사회의 각종 분야로부터 더욱 관심을 받고 있고, 빅데이터를 효율적으로 수집, 분석, 활용하려는 다양한 움직임이 있다.

국가적으로는 빅데이터 관련 추진정책이 발표되었고¹⁾ 공공에서 민간으로 빅데이터를 활용한 빅데이터 서비스가 확대되었다.²⁾ 도서관 분야에서도 도서관 이용자 특성, 장서구성, 대출현황, 웹컨텐츠 이용률 등 도서관의 다양한 데이터를 체계적으로 수집, 과학적으로 분석하여 도서관 서비스에 반영할 필요성이 대두되었다(심원식, 배경재, 박희진, 서유진, 안혜연, 서규원, 2015).

현재 공공도서관은 기존의 도서관리시스템인 KOLAS나 ILUS 등으로 데이터를 분석해 오고 있다. 이 시스템들은 회원 수, 대출이용자

수, 자료실별 소장자료 수, 주제별 장서 수, 대출건 수 등을 위한 일일통계, 매월 말 통계, 연도별 통계와 같은 단순통계를 제공한다. 하지만, 조회 조건이나 조합할 수 있는 내용이 단순하기 때문에 도서대출 현황 분석, 이용자 대출 분석 및 장서 분석과 같은 고급 분석은 조회된 수치에 대해 수작업으로 자료를 만들어 사용하며 그래프나 도표의 경우 별도로 공을 들여 작업해야 하는 등 통계작업 시간이 많이 소요된다. 더욱이, 변화하는 이용자 요구에 맞춤형 정보서비스를 제공하기에는 여러모로 한계에 부딪히게 되었다.

이러한 현재의 한계를 극복하고, 과학적인 분석 자료를 기초로 한 데이터를 확보하기 위해 문화체육관광부와 한국과학기술정보연구원(KISTI)은 2014년부터 빅데이터를 활용한 ‘도서관 빅데이터 분석활용 체계 구축사업’을 추진하였다. 이는 도서관 빅데이터 플랫폼의 분석 가능한 데이터를 활용함으로써, 이용자 정보 요구를 파악한 진일보한 서비스를 할 수 있는 계기가 되었다. 하지만, 도서관 빅데이터 플랫폼 활용사례가 아직은 초기단계인 만큼 도서관 빅데이터의 플랫폼 구축 및 활용에 대한 지속적인 데이터 발굴이 필요하며, 향후 추가적인 이용자 서비스 요구의 확장 가능성을 열어두고 도서관 빅데이터 플랫폼의 한계점이나 새로운 활용방안을 제시하여 다양한 도서관 빅데이터의 활용을 고도화해야 할 과제가 여전히 남아있다.

1) 우리나라는 2018년 6월 ‘데이터 산업 활성화 전략’, 2019년 1월 ‘데이터·AI경제 활성화 계획’(데이터와 AI의 통합적 전략 수립)을 발표해 세계적 데이터 경제흐름에 대응하고 있다.

<<https://www.4th-ir.go.kr/article/download/228>>

2) 공공빅데이터를 민간영역에서 활용할 수 있도록 하기 위하여 2013년 「공공데이터 제공 및 이용 활성화에 관한 법률」이 제정되어 시행되어오고 있다. <<http://www.law.go.kr/법령/공공데이터의제공및이용활성화에관한법률>>

이에 본 논문은 2016년 1월 1일부터 서비스하기 시작한 도서관 빅데이터 플랫폼 활용사례들을 살펴보고 도서관 빅데이터의 활용효과를 높일 수 있는 개선방안을 도출하고자 한다. 연구목적을 이루기 위한 방법으로, 첫째, 문화체육관광부와 한국과학기술정보연구원(KISTI)에서 2016년 1월 1일부터 서비스하기 시작한 도서관 빅데이터 플랫폼을 활용한 사례도서관들을 중심으로 도서관 빅데이터 플랫폼의 활용 데이터와 분석내용, 기능구현 등을 통해 도출된 서비스나 시행정책에 대해 고찰한다. 둘째, 현재 대전광역시 한밭도서관의 통합도서관정보시스템이 제공하는 데이터 분석사례와 도서관 빅데이터 플랫폼의 빅데이터를 활용한 분석 사례를 비교 분석한다. 이를 통해 도서관 빅데이터 플랫폼 활용 데이터분석의 특징과 한계점을 도출한다. 이를 토대로 공공도서관 이용자 정보서비스 향상을 위한 효과적인 도서관 빅데이터 플랫폼 활용방안을 제시하고자 한다.

2. 이론적 배경

2.1 도서관 빅데이터

빅데이터를 활용하여 가치와 유용한 결과를 도출하고자 하는 시도는 산업, 과학기술 및 학술 정보 영역뿐만 아니라 다양한 영역에서 이슈화되고 있다. 최근, 빅데이터가 이슈화되면서 도서관에서도 빅데이터에 대한 관심은 물론 도서관 빅데이터를 분석 및 활용한 업무의 효율성 향상 및 이용자 정보서비스 개발에 힘쓰고 있다.

도서관 빅데이터는 원시데이터를 만들어내

는 주체에 따라 다음의 4가지를 고려하여 수집이 가능하다. 첫째, 사서들이 만들어내는 원시 데이터이다. 수서작업, 목록작업 등과 같이 도서관 업무에서 생성되는 데이터이다. 둘째, 이용자들이 만들어내는 원시데이터이다. 대출, 검색, 참고서비스, 소셜미디어, 프로그램 참여 등을 통해 생성되는 데이터이다. 셋째, 도서관이 소장하고 있는 장서 그 자체의 콘텐츠 데이터이다. 학술논문, 특허, 보고서 등의 원문 텍스트와 초록, 목차 등과 같은 서지 텍스트가 이에 해당된다. 넷째, 도서관 외부의 사회적 데이터이다. 주로 공공 데이터 형식으로 공개되는 데이터를 많이 활용할 수 있을 것이다(이수상, 2014).

도서관 빅데이터는 기존의 장서, 회원, 통계 데이터뿐만 아니라 도서관 내·외부에서 생성되는 로그정보, 텍스트 정보 등이 모두 포함된다. 이러한 정형데이터, 반정형데이터, 비정형 데이터들은 수집과 분석, 적절한 활용을 통해 그 중요성과 가치가 비로소 분명해진다.

도서관 빅데이터 서비스의 가치는 1) 도서관 업무의 효율성을 가져다준다. 추천도서 제시, 자동화된 알고리즘으로 의사결정지원, 프로그램에 대한 운영정책의 선정 근거 확보 등으로 업무정책 방향 설정 등에 도움을 주어 업무시간을 절약하고 업무의 효율성에서 오는 비용 절감 효과를 가져다준다. 2) 도서관 이용자의 대출분석, 봉사대상 지역특성 등 이용자의 정확한 데이터 분석으로 이용자의 요구에 부응하는 정보 서비스를 발굴하고 도서관 품질을 향상하여 공공의 이용자 편익을 위한 서비스 제공에 있다(표순희, 김윤희, 김혜선, 김완중, 2015).

도서관 빅데이터는 도서관 업무의 효율성과 이용자의 정보서비스 만족도 향상이라는 두 가

지 핵심가치를 두고, 사서들이 만들어내는 원시 데이터, 이용자 정보요구 행태에 기초한 이용자 데이터, 소장 콘텐츠, 외부데이터 등 각종 데이터를 다양한 경로를 통하여 수집하고 분석하여 업무정책 방향을 설정하고, 신뢰성과 정확성을 담보로 한 양질의 정보서비스를 제공하여 이용자가 도서관 정보서비스에 만족할 때 도서관 빅데이터의 가치가 발휘될 수 있을 것이다.

이와 같이 다양한 경로로 수집된 데이터의 활용이 성공하기 위해서는 가치 있는 양질의 데이터 수집이 중요하다. 데이터 품질이 낮으면 좋은 정보, 가치 있는 정보를 얻기가 어렵다. 이와 같이 빅데이터의 공유와 연계를 통한 효과를 높이기 위해서는 빅데이터 플랫폼의 구축이 필요하다. 빅데이터 플랫폼 구축을 통해서 원하는 유용한 정보를 도출하고 빅데이터 분석기술을 통해서 효과적이고 체계적인 데이터 관리와 빅데이터를 활용하여 이용자 정보서비스 개발이 앞으로는 더욱 중요해질 것으로 예상된다.

2.2 선행연구

2013년 이후 도서관 빅데이터 활용에 대한 연구가 활발히 이루어져 오고 있다. 이정미(2013)는 빅데이터 개념의 재정립, 빅데이터 관련 이슈를 문헌과 사례를 통해 분석함으로써 빅데이터의 도서관정보서비스에의 활용 가능성을 밝혔다. 조영임(2013)은 지역정보화 차원에서 빅데이터 연구 시 제기되거나 고려되어야 할 이슈들을 논하였다. 김선태(2013)는 빅데이터시대에는 전문도서관 사서들에게 데이터과학자와 데이터큐레이터로서의 역할이 요구된다고 주장하였다.

표순희(2015)는 공공도서관에서 빅데이터 서

비스를 활용할 경우 얻게 될 경제적 가치를, 도서관구입 시 선정시간 절감효과와 이용(대출)증가에 따른 생산성 증대효과로부터 측정하였다. 표순희 외(2015)는 도서관 빅데이터 서비스에 대한 수요자 요구분석을 통하여 서비스 모형을 도출하였다. 이혜진, 김완중, 김혜선(2015)은 도서관 빅데이터를 활용하기 위한 분석 플랫폼 구축사례를 중심으로 이 플랫폼의 서비스 방안을 제시하였다.

2016년에는 도서관 빅데이터 플랫폼이 구축되어 서비스를 시작한 시기로 문화체육관광부와 한국과학기술정보연구원(KISTI)은 “도서관 정보나루”, “솔로몬(사서의사결정지원시스템)”, “R기반 통계분석도구”라는 공공도서관 빅데이터 플랫폼을 구축하였다(김완중, 2016).

공공도서관 이외의 타관종도서관에서 빅데이터를 적용한 연구사례들도 있었다. 한중희, 권상일, 김현숙, 서영주, 김난규(2014)는 국회도서관이 지금까지 제공해오던 『한국 박사 및 석사 학위논문 총 목록』의 표제와 초록을 대상으로 빅데이터 분석기법을 적용하여 국회도서관 학위논문 서비스인 2차 분석 콘텐츠(가칭 “학위논문연감”) 개발을 제안했다. 심원식 외(2015)는 국립중앙도서관이 자체적으로 보유한 다양한 디지털 자료들을 데이터화한 디지털 데이터 컬렉션을 구축하고 이를 최대한 활용하도록 디지털 자료의 질을 높이고, 이용자 커뮤니티를 개발하고 관련된 IT 인프라를 구축할 필요성을 주장하였다. 김태영, 백지연, 오효정(2018)은 국립세종도서관의 도서관 이용자 로그 빅데이터를 분석하여 이용자의 연령별, 성별, 거주지별 이용자 현황분석 및 연도별, 월별, 요일별 대출현황 분석을 수행하였다. 또한, 이용자의 현황 요인 분석

을 위하여 FGI 및 외부데이터와의 연계분석을 수행하였고 결과로 세종도서관의 운영에 도움을 줄 수 있는 개선방안을 도출하였다.

이원상(2014)은 대학도서관에서 학술정보 영역의 보다 좋은 정보를 선별하고 제공하기 위한 필수요소로서의 빅데이터 플랫폼 및 활용사례를 중점적으로 다루었다. 김서(2013)는 KAIST 도서관의 웹로그분석시스템을 구축하여 2013년 1월부터 도서관 데이터를 수집, 분석하여 웹 서비스를 개선하고자 하였다. 로그분석활용 서비스 개선사례로는 첫째, 유입 URL 상세분석을 통하여 '카이스트 도서관'으로 노출도서관 명칭을 변경하였으며, 둘째, 시간대별/요일별 방문자 수 분석을 통하여 최소방문 시간대에 서비스를 중단, 이용자 불편을 최소화하고, 셋째, 불필요한 페이지 로딩 문제를 발견하고 해결하여 시스템 부하를 감소시켰으며, 인터페이스 클릭수를 분석하여 많이 이용하는 메뉴를 메인화면에 비치, 이용자 편의성을 향상하였다.

3. 공공도서관 빅데이터 플랫폼과 활용사례 분석

도서관 빅데이터 플랫폼 구축은 2014년 문화체육관광부 R&D로 시작하여 2017년부터 국립중앙도서관이 주관하고 KISTI가 사업을 수행하였고, 사서의사결정지원서비스(2014~), 도서관 정보나루(2015~), R기반 통계분석도구(2014~), 오픈 API(2016~)가 구축 서비스 되고 있다.

도서관 빅데이터 플랫폼이 구축된 목적은 첫째, 합리적인 장서개발 및 수서정책 수립이다. 유형별, 주제별 도서대출현황에 근거한 장서구성 및 개발지침 방향성을 확립하고, 대출빈도가 높은 장서목록을 파악하여 수서할 때 적용하고, 베스트셀러 등 인기대출도서를 실시간으로 파악하여 도서관 이용자 만족도를 제고하고자 하는데 있다. 둘째, 장서이용률 향상을 위한 서비스 개발이다. 한번도 대출되지 않았던 미대출 장서수와 미대출 장서 비율을 주제 분야별로 파악하여 미대출 장서의 활용을 위한 독서 이벤트 및 프로그램을 개발하는 것이다. 셋째, 이용자의 연령대별, 성별, 도서대출 패턴 등에 대한 이용행태를 분석 활용하여 도서관 주요핵심 이용자를 비롯한 잠재 이용자의 특성 파악이 가능하며 월별, 연령별, 주제별 최다 대출목록을 게시하여 홍보가 가능하다. 넷째로 도서관 이용자 맞춤형 서비스 제공이다. 이용자의 연령대별 도서관 이용률 분석을 통한 관심 장서 주제분야를 파악한 맞춤형 문화프로그램 개발 및 운영이 가능하다. 예로, 이용률, 도서대출률이 높은 요일에 관련 주제분야 저자특강을 개설하고, 이용률, 도서대출률이 낮은 요일에 이용자 유입을 위한 문화프로그램 운영 등으로 국민 독서 진흥 및 도서관 내부 의사 결정을 지원하는 유용한 가치창출을 위한 사업이다.

3.1 공공도서관 빅데이터 플랫폼 이해

3.1.1 도서관 정보나루³⁾

도서관정보나루는 도서관 빅데이터 활용에

3) 도서관 정보나루. <<http://data4library.kr/>>

관심 있는 연구자, 개발자, 도서관을 위해 다양한 도서관 데이터를 개방 및 공유하여 새로운 서비스, 연구 개발에 활용할 수 있도록 지원하기 위한 분석서비스이다. 주요기능으로는 수집 데이터 통계, 기간별/지역별/성별/연령별 베스트 대출 장서 등 전국 공공도서관 주요 현황을 제공하고 전국의 서지데이터, 대출 데이터, 대출도서 순위 등 다양한 유형의 도서관 빅데이터 및 분석데이터를 제공하고 있다.

〈그림 1〉은 도서관 정보나루 홈페이지의 인터페이스이다. 데이터 수집현황을 가독성이 좋은 숫자와 그래프로 배치해서 데이터 현황을 쉽게 파악할 수 있다. 유용한 데이터로 다양한

메타데이터가 있으며, 참여도서관 인기대출 도서가 지역별, 연령별, 성별로 제공되고 있으며 빅데이터 활용현황을 현재일자로 표시되고 전체적으로 이해하기 쉽게 시각화되어 있다.

다음은 〈표 1〉에서 도서관 정보나루 주요기능을 간단하게 정리해 보았다.

도서관 정보나루는 2020년 5월 20일 현재, 1,003개 공공도서관이 참여하여 수집된 장서목록과 이용회원자료를 메타데이터로 하여 대출현황을 중심으로 분석하여 정보를 제공하고 있다. 기본적으로 메인화면에 데이터 수집 현황과 지역별, 연령별, 성별 다대출 도서에 대한 분석 결과를 제공하고 있다.



〈그림 1〉 도서관 정보나루 인터페이스
(<http://data4library.kr/>)

〈표 1〉 도서관 정보나루 주요기능

메뉴	주요기능
스마트 대시보드	<ul style="list-style-type: none"> • 수집 데이터 통계 현황판(장서, 회원, 대출, 위치 등) • 지역별, 연령별, 성별 인기대출도서 분석 • 데이터 활용현황: 파일 다운로드 수, API 호출건수
공개 데이터	<ul style="list-style-type: none"> • 데이터 공개 참여도서관 현황 • 참여도서관의 장서목록 및 대출 빈도 공개 • 다대출도서 공개(기간, 성별, 연령, 주제 등)
테마 데이터	<ul style="list-style-type: none"> • 수집데이터 기반 데이터 분석, 시각화 이미지 및 소스데이터
데이터 활용	<ul style="list-style-type: none"> • 오픈API 활용방법(회원 가입 후 인증 키 발급) • 텍스트(CSV파일) 활용방법 • 사용자 매뉴얼 및 오픈API 매뉴얼 제공
정보나루 소개	<ul style="list-style-type: none"> • 도서관 정보나루 소개 • 도서관 빅데이터 사업 소개(R 분석 툴 Radar 제공)
고객지원	<ul style="list-style-type: none"> • 공지사항, 질의응답

현재 공개데이터로 다운로드할 수 있는 것은 1) 참여도서관의 장서대출목록을 제공하고 월별로 자동 갱신하여 데이터의 연속성과 최신성을 유지하고 있다. 2) 다대출 도서를 대출기간, 성별, 연령, 지역, KDC 분류주제 등으로 검색이 가능하다. 3) 흥미를 유발하는 다양한 데이터 분석 결과를 제공하고 있는데 테마는 예를 들어, 세대별 선호 사회과학도서, 책과 함께하는 연말연시, 40대 여성 스테디 대출도서, 맨부커상 수상작 대출현황, 신록의 계절 읽기 좋은 책, 세대별 대출현황, 5월 가족의 달, 서가에서 빛나는 가족이야기, 남자들이 좋아하는 책/여자들이 좋아하는 책 등 특정 이슈가 되는 내용과 같은 다양한 테마 데이터를 제공하고 있어서 책을 선정할 때 많은 도움을 주고 있다.

3.1.2 사서의사결정지원서비스⁴⁾

사서의사결정지원 서비스는 도서관 내부에서 수집 가능한 장서 데이터, 이용자 데이터, 대

출 데이터와 온라인 서점데이터, 공공 데이터 등 도서관 외부 데이터를 수집, 분석하여 도서관 운영을 위한 사서 업무를 지원하는 웹기반의 의사결정지원 서비스이다. 사서업무지원을 위해서 자관 특성에 적합한 수서 및 업무 지원, 타관과의 연계분석 지원이 가능하도록 하며, 좀 더 세부적으로는 도서관별 일반현황, 회원 분석, 장서분석 결과 제공, 기간별, 분류별 선택을 통한 〈표 2〉 사서 데이터 조회, 차트, 테이블 등을 이용한 분석 결과에 대한 시각화 기능을 개발하였다.

홈페이지 첫화면 〈그림 2〉를 살펴보면, 현재 현황으로 자관 회원 수, 장서 수, 대출 건수가 표식화 되어있으며, 최근 1개월 동안 최다 대출 회원, 최다 대출 장서, 연령별 최다 대출 장서 순위를 알려준다. 무엇보다도 숫자들이 눈에 잘 들어오는 그림, 분석그래프 등으로 시각화 기능이 제공되어 이용자들이 쉽게 데이터를 읽을 수 있게 되었다.

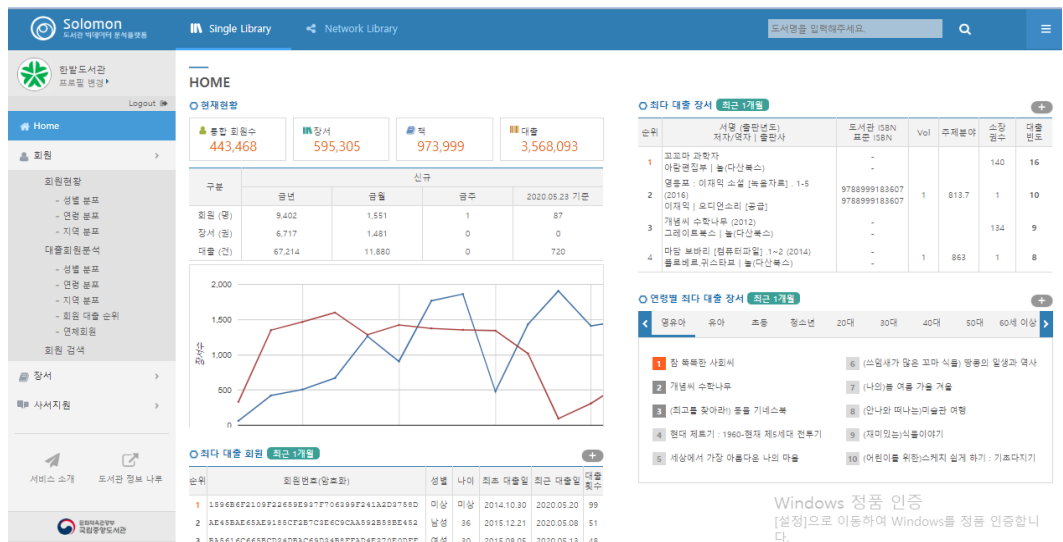
4) 사서의사결정지원서비스. 〈<http://solomon.data4library.kr/>〉

〈표 2〉 사서의사결정지원서비스 주요기능

메뉴	세부 디렉토리	분석 내용	
Home		<ul style="list-style-type: none"> • 현재 및 신규 데이터 현황 • 최근 1개월 최대 대출 장서 및 회원 현황 	
Single Library	회원	회원 현황	• 성별, 연령, 지역분포로 구분, 신규/누적회원수
		대출회원 분석	• 성별/연령별 대출·미대출회원, 1인당 평균대출, 지역분포도, 회원대출 순위, 연체회원, 자관/타관회원분석
	장서	회원 검색	• 다양한 조건으로 검색 가능, 회원의 대출 추이, 주제별 분포, 추천도서목록 제공
		장서 현황	• 연도별 장서분포, 기간별 주제분포, 동적차트 제공
		대출장서 분석	• 대출장서순위목록, 대출주제순위목록, 대출반납연체현황 주제별 비교 및 시간대, 일자별로 조회
		미대출장서 분석	• 우리 도서관의 미대출장서를 주제별로 조회 및 반출
장서 평가	• 장서별 대출반감기*, 주제별 평균 대출반감기*, 연도별 장서회전율*, 주제별 장서회전율* 조회		
장서 검색	• 우리 도서관 소장 장서를 다양한 조건으로 검색 • 서지, 책소개, 대출추이, 키워드 클라우드, 함께 대출된 장서 등 조회		
사서	사서 지원	• 자료 다운로드, 게시관 등	
Network Library	도서추천	<ul style="list-style-type: none"> • 알라딘 제공 베스트셀러 순위 및 신간도서 목록과 자관 보유현황 확인 • 전국도서관 기준 최근 3개월 다대출 순위와 자관 보유현황 확인 • 대출장서순위 목록 	
	외부데이터 연계	• 기상데이터 연계하여 기관별 자관대출 분석, 자관대출장서 순위, 전국대출장서 순위	
	관내 도서관 간 비교분석	• 연도별 신규/누적회원수, 대출회원 및 미대출회원 현황, 연체 회원 현황, 연도별 대출 및 연체회원, 자관/타관 대출 분석, 대출장서 비교분석 등	
	관내 도서관 통합분석	• 관내 도서관별 회원/장서/대출 현황 비교분석 요약 조회, 대출회원 순위 분석, 연체회원 분석	
	타 도서관과 자관 비교 분석	• 장서대출/주제별 수서 현황/장서회전율* 비교	
	전국단위 도서관 분석	• 전국 대출 분석, 연도별 대출분석, 주제별 대출 분석	
광역지자체별 도서관 분석	• 회원현황, 대출회원분석, 대출장서분석		

* 대출반감기란 대출건수가 1/2로 감소되기까지 소요되는 시간을 말한다.

* 장서회전율이란 도서 대출횟수를 소장 도서의 수로 나눈 값이며 장서회전율을 값이 클수록 분석 대상 도서의 대출횟수가 많음을 의미한다.



〈그림 2〉 사서의사결정지원서비스 인터페이스
(http://solomon.data4library.kr/)

메뉴에서 Single Library와 Network Library로 구분한 것을 볼 수 있는데, Single Library는 자관 소장데이터 및 이용데이터 기반 통계 분석자료이며, Network Library는 외부데이터(서점) 및 공공도서관 데이터와 비교하여 분석한 데이터를 표시하였다. 기존 도서관리시스템에서 제공하는 통계데이터와의 차별점은 알아보기 쉽게 그림, 도표 등으로 시각화하였으며, 외부데이터(타도서관, 서점 등)와의 비교 분석을 통해 우리 도서관 데이터의 현황을 편리하게 제공받을 수 있다는 점이다.

도서관 의사결정지원 서비스는 자관인 한발도서관이 소장하고 있는 장서, 한발도서관 회원 현황 등 정보를 수집하여 분석한 정보를 도서관 사서들에게 제공하여 도서관 운영과 관련된 의사결정을 지원하는 서비스이며 효율적인 정책 방향이나 결정에 활용 가능한 서비스이다. 예를 들면 성별, 연령, 지역 구분으로 연도별 신규 및 누적 회원 현황을 제공하고, 지역별 대출 회원 수, 대출 건수, 회원 1명당 평균 대출 건수를 파악하여 도서관 대출 회원을 분석할 수 있다. 또한 장서별, 주제별 대출빈도 순위를 파악하고, 기간별 대출주제 분포, 주제분야별 미대출장서 현황 확인, 소장 장서들에 대한 회전율, 대출 반감기 등을 실시간으로 분석, 평가함으로써 장서의 활용 정도를 파악할 수 있고 이를 통해 자관의 장서 구성 및 공간 구성을 최적화할 수 있다.

〈표 2〉는 솔로몬의 홈페이지에서 제공하는 주요기능을 표식화하여 나타냈다.

3.1.3 R기반 통계분석도구(Radar)

R기반 통계분석 지원 프로그램은 R과 Java

와의 통합을 통해 이용자가 한번에 PC에 다운로드 받아서 활용할 수 있는 GUI 기반 R 통계 분석 도구로서, 비전문가도 손쉽게 활용할 수 있는 통계분석 지원 프로그램이다. 통계분석 프로그램으로는 R을 활용함으로써 분석결과에 대한 검증 부담이 경감된다. 현재 약 100만 건까지의 데이터 분석이 가능하다. 로딩시간은 약 20분, 교차분석은 1초, T-test는 3초가 소요된다. 프로그램 설치 경로는 <http://data4library.kr>에서 하단의 분석 프로그램을 클릭하여 활용하면 된다. 서비스 내용은 도서관 운영 및 관련 연구 활동을 지원하는 통계서비스 제공으로 기술통계, 회귀분석, 교차분석, 분산분석(ANOVA), t-Test 등 핵심적인 통계기능 및 쉽게 사용이 가능한 GUI 기반 데이터 편집과 분석결과를 제공한다.

3.1.4 오픈 API 제공

추천도서 목록 및 공개데이터를 API로 공개하여 민간에서 누구나 해당 기능을 활용하여 서비스를 구현 및 활용할 수 있도록 개방하고 있다. 제공내용은 빅데이터 분석 플랫폼을 통해 데이터를 공개하는 도서관의 기본 정보를 제공하고, 전국 사업 참여도서관이 공개에 동의한 도서관별 장서 및 대출 데이터 정보를 제공하며, 전국 사업 참여도서관의 대출정보를 기반으로 한 인기 대출도서 순위 목록을 제공한다.

또한, 동시대출 도서 목록 및 추천 도서 목록을 조회하여 출력 가능하며, ISBN을 입력하여 해당 도서의 상세정보 및 대출 정보(최근 90일)를 제공하며, 서지 및 대출정보를 기반으로 분석한 핵심 키워드 정보를 제공하고 있다.

3.2 공공도서관 빅데이터 플랫폼 활용 사례 분석

문화체육관광부 도서관정책기획단은 도서관 빅데이터 공유 및 활용 플랫폼을 구축하고 도서관 데이터, 소셜데이터, 공공데이터 등 다양한 데이터를 수집하여 분석할 수 있는 플랫폼을 구축하고 빅데이터 수집과 연구를 위한 국내 협력 네트워크를 구성하였다.

또한 도서관 빅데이터 활용사례 경진대회,

도서관 빅데이터 플랫폼 활용 세미나, 도서관 빅데이터 전문가 양성을 위한 교육을 개최하여 도서관 빅데이터 활용도를 활성화하기 위한 노력을 기울였다.

도서관 빅데이터 플랫폼 구축사례는 공공도서관 협력 워크숍(2017.3.23.)에서 ‘도서관 빅데이터 분석활용 체계 구축사업 참여 안내’라는 주제로 발표한 자료를 중심으로 설명하고자 한다.⁵⁾ 다음의 <표 3>은 2016년 도서관 빅데이터 활용 경진대회에 참가한 사례도서관(부천 시립

<표 3> 도서관 빅데이터 플랫폼 활용사례

사례 도서관	플랫폼	활용데이터	분석내용/기능구현	도출서비스/시행정책
동대문구 답십리 도서관	도서관 정보 나무	도서대출 장서순위목록	- 전국도서관 대출장서 순위 분석 - 대상연령층 인기도서 유형 파악	- 어린이 인기도서 『강아지똥』 샌드아트공연
	솔로몬	미대출 장서 자료 (주제별)	- 주제별 순위 분석 - 대출도서 대상 평점 및 리뷰 분석	- 저자강연회 - 독서골든벨
		대출장서 분석	- 요일별 시간대별, 대출현황 분석 - 연령대별 자관/타관 최대대출 분석	- 과월호 잡지 나눔 - 연체자 해방데이 - 『우리가 사랑하는 도서』 목록전시
	회원통계	- 이용회원 분석 - 소장장서 및 이용현황 분석	- 『답십리도서관의 모든 것』 인포그래픽	
부천 시립 상동 도서관	도서관 정보 나무	전국도서관 인기도서 현황	- 이용자 성별, 연령별 대출현황 - 도서검색 및 희망도서 연계	- 오픈 API 활용:자관 홈페이지 연계로 전국 도서관 인기도서 정보제공 - 맞춤형 서비스 제공 - 소장도서검색연계로 이용편의제공
	솔로몬	대출회원 통계 (연령분포)	- 인구대비 도서관 회원비율 분석 - 인구대비 대출/미대출회원수 분석 - 전년대비 도서관 회원수증감 현황분석 - 전년대비 대출회원수증감 현황분석	- 도서관 회원 증가 현황 신속파악 - 북스타트대상연령 확대 - 금년도 추진정책 중간점검 - 내년도 데이터 기반 주요업무계획 수립
		대출/미대출 장서분석	- 미대출장서분석 - 연령별 관심도서분야 분석 - 대출장서 순위목록 분석 - 연령별 수요중심 분석	- ‘숨은도서찾기’ 코너 생성으로 대출 유도 - 홈페이지 사서추천도서 게시 - 수서참고자료로 활용
인천 서구 석남 도서관	솔로몬	대출회원 분석	- 이용자 연령, 성별 등 대출회원 분포 분석 - 이용자 대출 통계분석	- ‘우리도서관 예체능’ 운영테마 선정
		미대출 장서분석	- 연도별, 주제별 장서회전을 비교 분석 - 미대출장서 대상별, 주제별 분석	- 장서운영방향 및 계획수립에 활용 - 예술분야 테마서가, 기획도서전운영 - 예체능주제프로그램운영

5) 2017 공공도서관 협력 워크숍. (사)공공도서관협의회, 2017. <http://www.kpla.kr>

상동도서관, 동대문구 답십리도서관, 인천 서구 석남도서관)의 도서관 빅데이터 플랫폼 활용 데이터, 데이터 분석내용 및 기능구현을 통해 도출한 주요 서비스 영역과 시행정책을 정리한 것이다.

3.2.1 동대문구 답십리도서관

동대문구 답십리도서관은 도서관 정보나루의 활용데이터인 타도서관 최다대출도서 및 베스트셀러 데이터를 분석하여 정기도서, 신간도서, 복본도서 구입 등 수서정책에 적용하였다. 주제별 장서분석과 대출 분석을 참고하여 차년도 장서구성 계획 수립 시 활용하였다.

미대출장서를 활용데이터로 분석(주제별)하여 장서회전을 증가를 위해 책하나 더 읽기 '4+1' 행사, 책 속에서 보물찾기 등 다양한 행사를 진행하였다. 미대출도서를 활용하여 각 층 자료실별 테마서가를 조성하여 미대출장서를 활성화시켰다.

회원통계(성별, 연령, 지역분포), 대출회원 분석(성별, 연령, 지역분포 최다대출자 순위)을 통하여 도서관 이용자 현황 관리와, 최다 대출자 관리로 다독자상 시상 활용, 책읽는 가족 선정 시 활용, 도서관 우수 이용자 활용, 연체자 관리, 이용률이 저조한 연령대 타겟 홍보마케팅으로 활용하고, 최다 대출목록 게시 등 다양한 정보를 공유하였다. 또한 회원통계(성별, 연령, 지역분포), 회원별 대출주제분석, 대출장서 분석을 통해서 저자특강 및 테마특강 개최, 주제별 테마 서가 만들기, 생애주기별 맞춤형 프로그램 개발, 독서동아리 구성, 인문독서아카데미 개최 시 주제도서 선정에 활용하였다.

결국 도서관 빅데이터 플랫폼을 활용하여 데

이터 공유 및 활용을 통한 합리적 의사결정과 빅데이터 활용을 통한 선순환적 구조를 구축하고, 이용자를 위한 맞춤형 서비스 제공으로 선호 프로그램에 대한 선제적 대응과 이용자 만족도 및 도서관 이미지를 제고하였다.

3.2.2 부천시립 상동도서관

부천시립 상동도서관은 솔로몬의 대출장서 순위 목록을 분석하여 인기도서현황을 공개하고 리스트에서 도서 선택 시 도서검색서비스에서 해당 도서명 검색결과를 표출하였다.

또한 대출회원 분석을 통해서 인구대비 도서관 회원, 인구대비 대출회원, 미대출 회원, 인구대비 도서관 회원, 전년대비 대출자 수 등 이용현황을 분석하여 이용자 증감현황을 신속히 파악할 수 있었다. 미대출회원비율(32.7%)을 분석하여 열람실 회원제와 같은 회원유입전략을 모색하였으며, 영.유아 연령대의 이용률(2%) 부족과 관련하여 북스타트 대상연령을 확대함으로써 이용률을 향상시켰으며, 50세 이상 연령대 이용률(11%)의 증대를 위한 큰글코너, 책사랑회 구성 등 추가 프로그램 개발을 하였다.

미대출장서 분석(솔로몬)으로 '숨은 도서 찾기' 코너를 운영하여 운영 전에 비하여 미대출 도서 129권이 1회 이상 대출 발생하였으며, 사서추천도서는 1회 이상 대출발생 건이 제로에서 14건으로 증가하게 되었다. 또한 대출장서 순위 목록을 연령별로 수요중심 분석을 통해 수서참고자료로 활용하였다.

3.2.3 인천 서구 석남도서관

인천 서구 석남도서관은 솔로몬을 활용한 도

서관 대출회원을 분석하여 주 이용계층의 성별과 연령을 파악하고, 자체 설문조사를 통해 이용자의 수요 및 특성을 분석하였으며, 도서관 증점 활성화 분야 도출하여 '우리 도서관 예체능' 운영 테마를 선정하고 테마에 부합한 장서구성, 문화행사 및 프로그램을 운영하였다. 먼저, 솔로몬의 장서를 활용하여 장서운영의 기본방향 수립시 솔로몬의 장서 관련 데이터를 분석하고, 이를 바탕으로 장서 현황 파악 및 정책을 수립하고, 2016년도 장서 '우리 도서관 예체능'으로 예술분야 테마서재를 운영하고, 주 이용계층의 수요를 충족시키기 위해 예술분야 장서 확충하고 예술분야(취미, 스포츠 등) 기획도서전을 운영하였다. 또한, 도서관에 가입한 회원의 연령과 성별을 분석하여 특성을 파악하고, 관심분야 및 이용패턴을 분석하여 이용자 요구사항을 반영하여 취미·예술분야의 연령별 평생학습프로그램 개발하고, 유아·아동을 위한 오감발달 미술놀이 운영, 성인을 위한 금요공작소 운영 등 이용자의 요구를 반영한 평생학습프로그램을 개발하였다.

3.2.4 종합분석 및 개선방안

앞에서 살펴본 바와 같이 도서관 빅데이터 플랫폼 활용사례를 통해 나타난 활용서비스 및 성과 부분의 공통점을 도서관 업무 분야별로 <표 4>에 정리하였다.

사례 도서관들의 빅데이터 플랫폼 활용 서비스의 공통점은 운영프로그램 계획 근거, 장서운영의 활성화, 수서리스트 확보, 시각적인 통계자료 제공 등 대부분 도서관 업무관리의 효율적인 부분이다.

사례도서관들은 빅데이터를 이용해 대출회원의 성별 및 연령분포를 분석하고, 자관의 장서 회전율과 타 도서관의 장서 회전율을 비교하는 등 다각도로 데이터를 분석하여 도서관 운영방향을 정립하고 이용활성화를 위한 방안을 모색하여 다양한 문화프로그램 운영, 독서진흥행사 등을 수행하였다. 특히 인천 서구 석남도서관은 이용자 설문조사 및 정보요구 수요조사 결과에서 예술, 취미분야의 높은 관심도와는 반대로 빅데이터 분석결과에서는 장서보유량과 대출 회전율이 낮다는 점에 착안하여

<표 4> 사례도서관 빅데이터 플랫폼 활용서비스 공통점

구분	활용데이터	활용서비스	성과
프로그램 분야	대출회원통계 대출장서 순위목록	- 이용자 연령별 대출빈도 분석 - 이용자 대출도서 분석	- 업무정책방향 수립 기준 근거 - 강연 주제 및 프로그램 테마 결정에 과학적 분석 근거 마련
장서분야	미대출 장서 자료	- 미활용도서 분석 추천 게시 - 미활용도서 코너 마련 - 기획도서전(테마서가)	- 장서 활용도 증대 - 장서구성정책 근거
수서분야	대출장서 순위목록	- 베스트 대출도서 및 베스트셀러도서 활용 - 다대출도서순위 및 보유현황	- 신간도서 구입 선정에 유용 - 수서정책방향 설정 근거
기타	오픈API 제공 데이터	- 오픈API를 활용하여 전국 인기도서 및 테마도서 홈페이지 연계서비스 제공	- 이용자 도서정보서비스 제공 - 이용자가 쉽게 이해할 수 있는 시각적 효과에 도움

수요도와 대출률의 연계성을 높이기 위한 다양한 예술 프로그램 기획, 예술분야의 도서 전시 및 신간도서 확보 등 빅데이터 분석 결과에 맞추어 도서관 운영을 수행하였다. 그 결과, 이용자들로부터 긍정적 반응과 대출증가 효과를 가져왔다⁶⁾는 점에서 빅데이터 분석의 효과와 중요성을 재인식시키는 검증 결과를 보여 주었다.

도서관 빅데이터 플랫폼 활용사례로부터 얻어진 빅데이터 플랫폼의 개선방안을 제시해 보면, 첫째, 좋은 성공사례로 타도서관과의 비교 가능하다는 점이다. 이는 각 도서관의 현재 위치를 스스로 파악할 수 있도록 하여, 업무개선의 자발적인 동기부여와 도서관 중장기 발전방향을 가능하도록 하였다는 점이다. 현재 도서관 빅데이터 플랫폼은 대다수의 공공도서관으로부터 수집된 데이터를 분석하여 제공하고 있다. 이에 더 나아가 도서관 빅데이터 플랫폼이 빅데이터의 속성을 잘 반영하기 위해서는 타관종(대학도서관, 전문도서관, 학교도서관)의 데이터도 포함한다면 전체 도서관 데이터의 관점에서 개별 도서관의 빅데이터 플랫폼 활용이 효과적일 것이다.

둘째, 대출회원 통계 및 대출 장서 목록에 대한 현황에 대한 요인으로 도서관 내부 요인에 대해서만 가능하다는 한계가 있다. 도서관 외부요인에 의한 도서관 정보서비스 효과 및 품질의 영향은 미미한 실정이다. 예를 들어, 교통의 편의성과 같은 도서관 접근성 측면의 도서관 외부 요인(교통데이터)이 도서관 정보서비스 효과 및 품질에 미치는 영향, 인구이동(인구데이터)이 대출회원의 증감에 미치는 영

향, 도서관이 위치는 주변의 교육열(교육데이터)이 도서관 이용에 미치는 영향은 분석하기가 어려운 상황이다. 이는 더 많은 공공데이터와 도서관 빅데이터가 결합될 필요가 있음을 시사한다.

4. 통합도서관리시스템(ILUS)과 도서관 빅데이터 플랫폼 비교 분석

대전광역시 한밭도서관은 1989년 12월 20일 개관한 이래, 대전시민의 독서진흥과 양질의 정보서비스 제공뿐만 아니라 대전지역 도서관 통합정보관리시스템 구축 운영, 작은 도서관 운영자를 위한 세미나 개최, 공공도서관 대전지부 운영 등 대전광역시의 대표도서관의 역할을 수행하기 위하여 대전지역의 각종 도서관 지원 및 협력 사업을 추진하고 있다. 특히 통합도서관리시스템(ILUS)은 2015년부터 표준자료관리시스템(KOLAS III)에서 변경한 도서관리 프로그램으로 대전지역 23개 공공도서관의 산재한 데이터베이스 통합하였으며, 공공도서관과 작은도서관 연계시스템 기반을 마련하여 다양한 이용자 요구에 맞는 도서정보 서비스 제공과 업무 효율성을 증대시켜오고 있다.

본 장은 현재 한밭도서관에서 사용하고 있는 통합도서관리시스템(ILUS)과 도서관 빅데이터 플랫폼 제공 데이터와 비교분석함으로써 두 시스템의 특성들을 살펴보고, 도서관 빅데이터 플랫폼의 한계와 개선방안을 도출하고자 한다.

6) 한국과학기술정보연구원 정보서비스실 (2016). 도서관 빅데이터 활용사례집 [사례집]. 김완중: 편집인.

4.1 통합도서관리시스템(ILUS)

한밭도서관 통합도서관리시스템(ILUS)은 소프트웨어 구성도 <그림 3>에 보듯이 대전광역시 24개 도서관 업무 및 서비스 통합을 위한 통합솔루션 구축과 안정된 시스템 운영을 위한 시스템 이중화 및 백업 체계 구축에 주안점을 두고 구축한 시스템으로 데이터 분석을 위한 시스템 개발은 설계되어 있지 않다. 최근까지 한밭도서관 데이터 분석은 통합도서관리시스템(ILUS)을 기반으로 장서 및 회원 통계 중심의 자료를 활용하였다.

다음 <표 5>는 대전통합도서관리시스템에서 제공하는 통계데이터 중에서 조회 가능한 데이터(수치통계 중심)를 대출과 목록관리로 구분하여 정리하였다. 대전통합도서관정보시스템에서

제공하는 통계자료는 주제별 장서통계, 회원수, 이용자 도서대출건수를 중심으로 일별, 월별, 요일별, 분류별, 연도별, 자료실별로 산출하는 단순한 통계로 매일 작성되는 업무일지 통계 작업을 위해서는 유용하다.

4.2 ILUS와 도서관빅데이터 플랫폼 비교분석

도서관 빅데이터 플랫폼은 통합도서관리시스템에서 제공하는 자료실별(어린이자료실, 제1자료실, 제2자료실, 다문화자료실, 디지털자료실, 이동도서관 등) 통계와 신분별(일반인, 장애인, 다문화 다가족, 공무원으로 구분)로 대출하는 내역은 포함하고 있지 않다. 반면에 도서관 빅데이터(솔로몬 등)에서는 연령별/기간별/



<그림 3> 한밭도서관 통합도서관리시스템 구성도 (출처: 한밭도서관 연보, 2019)

〈표 5〉 대전통합도서관정보시스템 통계자료(기초분석자료)

대출관리	목록관리
최다대출도서 BEST	소장처/분류별통계
최다대출 이용자 BEST	자료실별언어통계
연체자 리스트	분류/유형별통계
회원가입자수 통계	주제별/유형별장서통계
이용자수, 대출건수 통계	분류통계
대출/반납통계: 일별, 월별, 요일별, 분류별, 시간별	소장처자료실별/유형통계
*신분별 대출반납 통계: 일별, 월별, 요일별, 시간별	언어별/유형별통계
자료실별 대출반납통계	자료실류별장서통계
상호대차 건수 통계	언어별분류통계
대출반납통계(자료실별)	
*자료실별 등급별 대출자수 통계	
자료실별 등급별 반납자수 통계	
자료실별/신분별 대출통계	
대출자수 통계	
*소장처별 장서수 통계	
소장처별 회원수 통계	
소장처별 대출/반납 통계	
소장처별 이용자수/대출건수 통계	

* 신분별: 일반, 장애인, 다문화, 다가족, 공무원으로 구분
 등급별: 어린이, 청소년, 성인, 노인으로 구분
 소장처별: 대전관내 24개 공공도서관으로 구분

성별/지역별로 세분화하여 대출데이터를 분석하여 제공하고 있다. 〈표 6〉은 ILUS와 도서관 빅데이터 플랫폼을 비교하여 두 시스템 간의 상관성과 상이성을 보여준다. 이는 도서관 빅데이터 플랫폼의 확장된 분석서비스를 이해하는 데 도움을 준다.

도서관 빅데이터 플랫폼은 기존에 제공하였던 수치 통계뿐만 아니라 수작업 없이 데이터 결과를 파악할 수 있으며 자동 분석된 결과를 통해 통계작업이 수월하다. 또한 도표, 그래프 등을 활용하여 데이터 분석이 훨씬 용이해졌다. 또한 외부데이터를 통하여 비교 분석이 가능해졌으며, 서점 등의 베스트 셀러 순위를 바로 파악할 수 있어서 수서업무가 수월하고, 오픈 API

활용을 통한 데이터 연계서비스 제공으로 이용자 도서 시각화, 도서추천 서비스 등 이용자 정보 서비스 만족도를 향상시킬 수 있다.

이상의 내용을 정리하면, 통합도서관리시스템(ILUS)에서 제공하는 통계자료는 대출건수, 회원수 등 단순한 기술통계이며 데이터를 분석하기에는 한계가 있으며, 고급통계를 사용하거나 과학적 기반의 분석으로 의사결정하기에는 역부족이다. 그에 비해 도서관 빅데이터 플랫폼은 실시간 현황자료 제공이 가능하고, 관내 도서관 간 비교분석 및 통합분석이 가능하고 외부데이터와 연계하여 데이터 분석이 가능하며, 대출빈도를 연령대도 세분화 되어 있으며, 성별/연령별 성별/지역, 연령/지역에 대한 교

〈표 6〉 통합도서관리시스템(ILUS)과 도서관 빅데이터 플랫폼과의 비교

[통합도서관리시스템]	[도서관 빅데이터 플랫폼]	
대출 및 목록 관리	디렉토리	분석 내용
최다대출도서 BEST 최다대출 이용자 BEST	현황판	- 현재 및 신규 데이터 현황 - 최근 1개월 최대 대출 장서 및 회원 현황
회원통계 - 연체자 리스트 - 회원가입자수 통계	회원통계	- 성별, 연령, 지역분포로 구분, 신규/누적회원수
대출분석 - 이용자수, 대출건수 통계 - 대출/반납통계: 일별, 월별, 요일별, 분류별, 시간별 - 신분별 대출반납통계: 일별, 월별, 요일별, 시간별 - 자료실별 대출반납통계 - 도서 대출자수	대출회원 분석	- 성별/연령별 대출·미대출회원, 1인당 평균대출, 지역분포도, 회원대출 순위, 연체회원, 자관/타관회원분석
자료실별(13개 자료실) - 등급별 대출자수 통계 - 등급별 반납자수 통계 - 신분별 대출통계 - 언어통계 - 류별장서통계 - 분류/유형별통계 - 주제별/유형별장서통계	회원검색	- 다양한 조건으로 검색 가능, 회원의 대출 추이, 주제별 분포, 추천도서목록 제공
소장처별(24개 공공도서관) - 장서수 통계 - 회원수 통계 - 대출/반납 통계 - 이용자수/대출건수 통계 - 분류별통계 - 자료실별/유형통계	장서통계	- 연도별 장서분포, 기간별 주제분포, 동적차트 제공
언어별/유형별통계 언어별분류통계 상호대차 건수 통계	대출장서 분석	- 대출장서 순위목록, 대출주제 순위목록, 대출반납 연체현황 주제별 비교 및 시간대, 일자별로 조회
	미대출 장서분석	- 우리 도서관의 미대출장서를 주제별로 조회 및 반출
	장서평가	- 장서별 대출반감기, 주제별 평균 대출반감기, 연도별 장서회전율, 주제별 장서회전율 조회
	장서검색	- 우리 도서관 소장 장서를 다양한 조건으로 검색 - 서지, 책소개, 대출추이, 키워드 클라우드, 함께 대출된 장서 등 조회
	사서지원	- 자료 다운로드, 게시판 등
	도서추천	- 알라딘 제공 베스트셀러 순위 및 신간도서 목록과 자관 보유현황 확인 - 전국도서관 기준 최근 3개월 대대출 순위와 자관 보유현황 확인 대출장서 순위 목록
	외부데이터 연계	- 기상데이터 연계하여 기관별 자관대출 분석, 자관대출장서 순위, 전국 대출장서 순위
	관내 도서관간 비교분석	- 연도별 신규/누적회원수, 대출회원 및 미대출회원 현황, 연체 회원 현황, 연도별 대출 및 연체회원, 자관/타관 대출 분석, 대출장서 비교분석 등
	관내 도서관 통합분석	- 관내 도서관별 회원/장서/대출 현황 비교분석 요약 조회, 대출회원 순위 분석, 연체회원 분석
	타도서관과 자관 비교분석	- 장서대출/주제별 수서 현황/ 장서회전율 비교

차분식이 가능하고 분석된 자료는 도표와 그래프로 시각화하여 쉽게 이해할 수 있도록 데이터를 제공하고 있으며, 무엇보다 회귀분석을 통한 예측이 가능하고 이러한 분석 자료를 기반으로 이용자 서비스 방향 수립이 가능하다는 것이 도서관계에서는 매우 고무적이다.

또한 R기반 통계분석 지원프로그램을 통해 기술통계, 회귀분석, 교차분석 등이 가능한 GUI 기반의 데이터 편집과 분석이 가능하며, 오픈

API 제공으로 홈페이지와 연계하여 데이터를 공개하고 전국 공공도서관의 대출정보를 제공할 수 있다.

4.3 비교 예시

4.3.1 대출통계 비교

다음은 한밭도서관 데이터 조회 및 분석 시에 실제로 어떠한 결과 값으로 나타나는지를 통합

도서관리시스템(ILUS) 자료실별/등급별 대출 통계(〈표 7〉 참조)와 도서관 빅데이터 플랫폼에서 조회하여 분석한 결과(〈표 8〉 참조)를 가지고 비교해 보았다. 통합도서관리시스템(ILUS)에서 제공하는 자료실별/등급별 대출통계(〈표 7〉 참조)는 자관

〈표 7〉 통합도서관리시스템조회 한밭도서관 자료실별/ 등급별 대출통계 예시

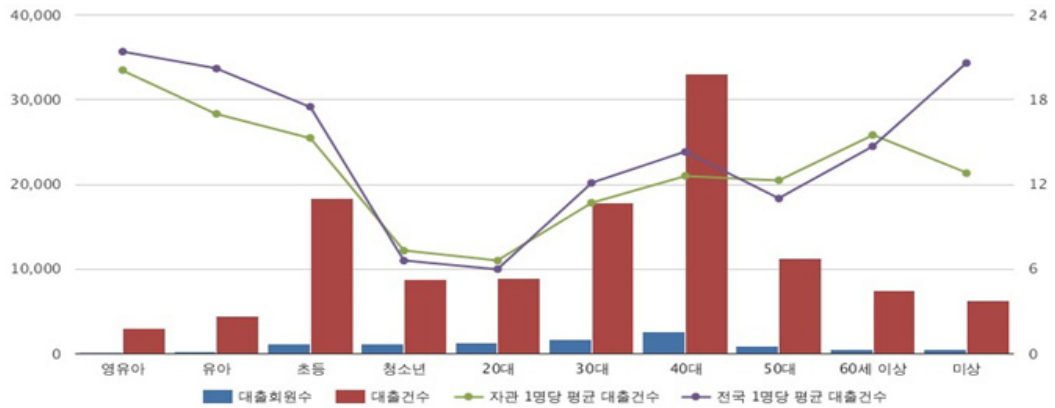
자료실별/등급별 대출 통계

검색조건: 소장처 = 한밭도서관 < AND > 대출일: 2016.01.01. ~ 2016.12.31.

자료실	어린이		청소년		성인		노인	
	남	여	남	여	남	여	남	여
1서고(지상1층)	38	57	146	154	1127	1258	321	108
2서고(지하1층)	715	784	80	91	828	1158	29	63
3서고(지하2층)	49	48	185	160	2049	1681	262	88
디지털정보센터	244	355	252	264	1692	1991	27	24
시각장애인실	1	0	3	4	20	53	1	1
어린이자료실	33558	36465	3191	3179	27034	42290	1192	2424
운영과사무실	0	0	0	0	0	0	0	0
이동도서관	675	630	32	88	1021	3318	99	81
정기간행물실	0	0	0	0	0	0	0	0
제1자료실	2460	2763	10220	13148	69738	78769	6346	2827
제2자료실	0	0	1	3	120	77	0	0
제3자료실	29	186	586	353	1050	1061	25	214
제적자료	9	14	3	3	8	7	1	0
청록실	0	0	0	0	0	0	0	0
폐기대상자료	100	68	48	34	117	150	4	4
하늘도서관	1	0	0	1	1	1	0	0

〈표 8〉 도서관 빅데이터 플랫폼(솔로몬) 조회한 한밭도서관 대출통계(2016.1.1.~12.31.) 예시

구분	대출회원수	대출건수	회원 1명당 평균 대출건수	
			자관	전국
영유아	150 (1.5%)	3,014 (2.5%)	20.1	21.4
유아	260 (2.5%)	4,423 (3.7%)	17	20.2
초등	1,200 (11.6%)	18,312 (15.3%)	15.3	17.5
청소년	1,202 (11.6%)	8,752 (7.3%)	7.3	6.6
20대	1,346 (13%)	8,925 (7.5%)	6.6	6
30대	1,678 (16.2%)	17,885 (15%)	10.7	12.1
40대	2,618 (25.3%)	33,059 (27.7%)	12.6	14.3
50대	919 (8.9%)	11,295 (9.5%)	12.3	11
60세 이상	483 (4.7%)	7,464 (6.3%)	15.5	14.7
미상	487 (4.7%)	6,253 (5.2%)	12.8	20.6
합계	10,343	119,382	11.5	12.4



〈그림 4〉 〈표 8〉의 분석결과와 함께 자동 생성되는 그래프 예시

실정에 맞도록 자료실 중심 통계이며 더 이상의 확장된 서비스는 제공되지 않는다. 도서관 빅데이터 플랫폼(솔로몬)에서 조회한 대출통계(〈표 8〉 참조)는 연령별뿐만 아니라 성별, 기간별, 지역별 등 다각도로 조회가 가능하며, 이용자가 이해하기 쉽게 그래프가 자동으로 생성되어 시각적인 정보 편의를 제공하고 있다.

장서수 현황이 조회되지 않아서 별도로 주제별로 검색하여 수치통계를 수작업으로 작성하여 만든 것을 나타냈다. 통합도서관리시스템은 도서관 빅데이터 플랫폼에 비해 주제별 장서수 현황 파악에 시간과 노력이 소모되는 반면, 〈표 10〉의 도서관 빅데이터 플랫폼(솔로몬) 데이터 서비스는 실시간으로 주제별 장서 현황 파악은 물론 자동생성되는 그래프가 제공되어 업무 이용에 매우 편리하다.

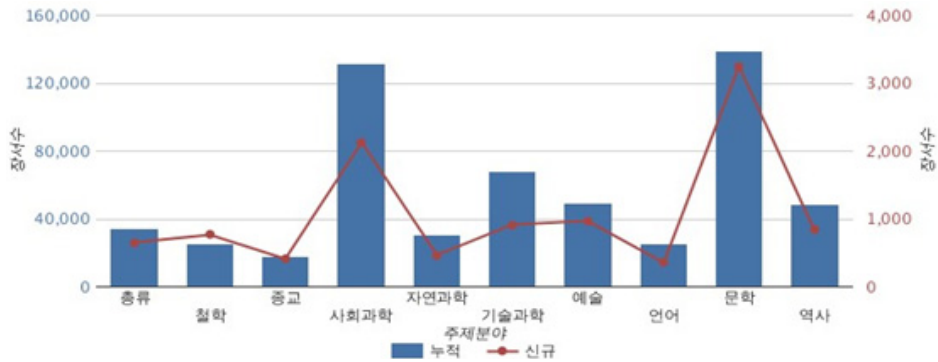
4.3.2 주제별 장서수 현황 통계 분석 비교
〈표 9〉는 통합도서관리시스템에서 주제별

〈표 9〉 통합도서관리시스템 데이터 조회후, 수작업한 주제별 장서수 현황 예시(2016년도)
(단위: 권)

주제별 구분	계	000 (총류)	100 (철학)	200 (종교)	300 (사회과학)	400 (순수과학)	500 (기술과학)	600 (예술)	700 (언어)	800 (문학)	900 (역사)
누계	802,933	48,376	41,308	28,877	185,639	45,178	95,825	45,193	28,778	218,255	65,504

〈표 10〉 도서관 빅데이터 플랫폼(솔로몬) 조회한 주제별 장서수 현황 예시: 2016.1.1.~12.31.

구분	총류	철학	종교	사회 과학	자연과학	기술과학	예술	언어	문학	역사	합계
신규	645 (6%)	774 (7.2%)	414 (3.9%)	2,129 (19.8%)	460 (4.3%)	907 (8.4%)	966 (9%)	360 (3.4%)	3,244 (30.2%)	843 (7.8%)	10,742
누적	34,168 (6%)	25,362 (4.4%)	17,618 (3.1%)	131,348 (23%)	30,728 (5.4%)	67,778 (11.9%)	49,537 (8.7%)	25,735 (4.5%)	138,733 (24.3%)	48,945 (8.6%)	569,952



〈그림 5〉 〈표 10〉의 분석결과와 함께 자동 생성되는 그래프

4.4 플랫폼 개선 방향

대전 통합도서관리시스템의 단순한 기술통계 제공에 더하여 도서관 빅데이터 플랫폼 구축으로 인한 데이터 분석은 타 도서관 사례에서 알 수 있듯이 많은 부분에서 도서관 업무의 효율성을 제고하고 이용자 서비스를 향상할 수 있는 데이터 분석이 가능해졌다. 도서관 빅데이터 분석을 통해 바로 적용할 수 있는 도서관 업무를 제시하면, 1) 주제별 장서현황과 이용자 대출현황을 통한 장서구입방향을 설정할 수 있다. 2) 외부데이터인 베스트셀러도서 및 공공도서관 최다대출도서 분석을 활용하여 신간도서 수서 시 반영할 수 있다. 3) 미대출 장서 분석을 통한 장서활용 방안을 강구할 수 있다. 4) 연령대별로 세분화된 대출장서 분석을 통해서 이용자 계층별로 맞춤형 서비스 제공 및 프로그램 개발할 수 있는 근거의 유용한 자료로 활용될 수 있다.

이와 같이 도서관 빅데이터 플랫폼 기반 데이터 활용은 공공도서관 자료 분석의 다양화, 시

간과 노력의 투자비용 감소, 통계작업의 간소화 등 커다란 변화를 가져오며, 도서관 데이터에 대한 과학적인 분석으로 업무의 효율성을 향상시킬 수 있다.

기존 ILUS와의 비교를 통해 얻어진 빅데이터 플랫폼 개선방안을 제시해 보면, 첫째, 현재 ILUS는 장서의 소장위치(1,2서고)나 특정이용자유형(어린이자료실, 청소년자료실, 종합자료실), 자료 유형(디지털자료실, 정기간행물실)에 따라 통계자료 분석이 가능하다는 장점이 있다(〈표 7〉 참조). 매일의 통계보고 및 의사결정이 자료실별 업무단위로 이루어지는 것을 감안한다면, 도서관 빅데이터 플랫폼도 이러한 개별도서관의 특성에 따른 자료실별이나 통계작성 및 분석을 지원하는 맞춤형 플랫폼으로 발전할 필요가 있다. 둘째, 도서관 빅데이터 플랫폼은 분석데이터의 신뢰성을 보장할 필요가 있다. ILUS시스템과 도서관빅데이터플랫폼의 최다대출장서목록을 비교분석하는 과정 중에, 도서관빅데이터 플랫폼에 오류데이터가 포함된 것을 발견하였다.⁷⁾ ILUS의 원본데이터는 오류

7) [부록 1]의 연령별/세대별 대출장서순위 목록 중 0-5세 대출내역 중 서명에서 영유아가 읽을 수 없는 책이 포함되었으며, 6-7세 대출내역에서도 “우리는 공산당이 싫어요”라는 책이 1위였으며, 이 도서는 한밭도서관도, 인터

가 없었으나 도서관빅데이터 플랫폼의 데이터에 오류가 포함되어 있었다. 확인결과, 데이터 업로드 과정에서 문제가 있었다. 이는 데이터의 입수에서부터 저장, 처리 및 분석, 배포에 이르기까지 신뢰성을 보장할 수 있는 데이터 검증과정이 필요하다는 점을 시사한다.

5. 도서관 빅데이터 분석의 한계점 및 개선 방안

앞에서 살펴본 사례분석과 비교분석을 통해 도서관 빅데이터 플랫폼의 서비스의 특징과 한계점을 아래와 같이 도출하였다.

첫째, 도서관 정보나루에서 제공하는 공공도서관 빅데이터는 1,003개관(2020.5.20. 현재)이 참여하여 수집한 데이터이다. 이 수치는 국가도서관통계시스템⁸⁾에 등록된 2018년 기준 전국공공도서관수인 1096개관(지자체 설립 840, 교육청 설립 233, 사립 23)을 고려할 때, 전체 공공도서관의 91.5%의 공공도서관으로부터 수집된 데이터이다. 이는 전체 공공도서관의 특성을 반영하는 대표성 있는 빅데이터라는 특성을 가지고 있다고 할 수 있다. 이에 더하여 빅데이터의 범위를 대학도서관, 연구도서관, 학교도서관 데이터로 수집을 확대한다면, 빅데이터의 속

성인 용량(volume) 및 다양성(variety) 측면에서 도서관 빅데이터로서의 충분한 가치를 만들어 낼 수 있을 것이다.

둘째, 도서관 빅데이터 플랫폼과 ILUS의 비교연구에서 밝혔듯이, 도서관 빅데이터 플랫폼은 여러 가지 유용한 데이터를 제공하지만, 아직은 현장에서 활용하기에는 자관별 통계 자료 분석 양상이 미반영되어 통계활용의 한계성을 가지고 있다. 예를 들어, 공공도서관 현장에서는 자료실별로 업무/정책방향 및 계획 수립, 시행, 성과평가가 이루어지고, 이를 위해 자료실별로 통계일지 및 업무일지 등을 작성하고 있다. 하지만, 도서관 빅데이터 플랫폼은 개별 도서관의 특성이 반영되어 있지 않고, 현장에서 자주 사용하고 있는 자료실별로 데이터 제공이 되지 않아 도서관 빅데이터 업무와 자관의 통계업무에 가중화가 초래되고 있다. 개별도서관자료(자료실별, 테마서가) 및 회원(회원 분류)의 범주화와 연관하여 다룰 수 있는 기능을 빅데이터 플랫폼에서 제공할 수 있도록 개선해야 한다.

셋째, 도서관 빅데이터 플랫폼에서 제공하는 데이터 분석 자료만으로는 다각적으로 원하는 분석결과를 얻기가 힘든 한계성을 가지고 있다. 이를 위해, 외부데이터와(인구, 기상, 교통 등)의 연계를 확대해야 한다. 또한 분석과정에서 Radar분석 틀을 사용하는 과정에서 고급통계

넷 서점도 보유하고 있지 않은 도서관이다. “행태주의 마케팅”이란 책 또한 저자명과 서명이 일치하지 않고 있었다. 14-19세 대출내역에서도 “전환기 중국의 정치경제”, “김밥 주먹밥 찜밥”이라는 도서가 대출순위 상위 10위권에 포함되어 있었으며, “전환기 중국의 정치경제”는 주제가 사회과학임에도 문학류로 표시되어 있는 등 분류기호 또한 오류를 포함하였다. 결국 도서관 빅데이터 플랫폼에서 제공하는 한밭도서관 대출 장서 순위 목록을 포함한 장서현황분석은 오류 데이터로 이용자의 자료대출 패턴파악이 어려웠다. 이후 오류내용은 한국과학기술연구원 도서관 빅데이터 담당자를 통해 정정되었다. [부록 2]는 정정된 결과를 보여준다. 원인은 한밭도서관 장서데이터의 업로드과정에서 문제가 있었다. 이러한 데이터 신뢰성 문제는 빅데이터 플랫폼의 분석데이터와 자관 데이터의 정합성 검증프로세스의 필요성을 제기한다고 하겠다.

8) 국가도서관간통계시스템. <<http://www.libsta.go.kr>>

분석에 대한 기본지식이 없는 경우 분석의 어려움을 겪을 수 있다. 도서관 빅데이터 플랫폼 적용 시 현장사서가 조금 더 쉽게 분석할 수 있는 시스템 또는 분석도구 개발 뿐만 아니라 또는 분석 툴 사용에 대한 지속적인 교육과 실습 기회가 제공되어야 할 것이다.

넷째로, 도서관 빅데이터의 가장 중요한 데이터 신뢰성의 검증시스템 관리에 대한 한계이다. 도서관 빅데이터 플랫폼 활용과정에서 장서현황 분석 데이터의 오류를 발견하였다. 이는 장서데이터의 업로드 과정에서 문제가 있을 수 있음을 보여준다. 이러한 일이 발생되지 않도록 도서관 빅데이터 관심 제고를 위한 다각적인 노력이 필요하다. 소장 서지데이터의 신뢰성을 높이고 수집데이터의 품질 제고를 위해 데이터 가공/처리 시 데이터 관리자가 오류데이터 여부를 바로 확인하여 정확한 데이터 검증할 수 있도록 시스템화하여 관리해야 할 것이다.

6. 결론

본 논문에서는 도서관 빅데이터 플랫폼을 일선 공공도서관 업무에 효과적으로 활용할 수 있는 방안을 도출하기 위하여, 2016년 1월 1일부터 서비스하기 시작한 공공도서관 빅데이터 플랫폼을 활용한 사례도서관의 빅데이터 활용 유형분석 및 시행정책을 살펴보고 현재 통합도서관리시스템과 도서관 빅데이터 플랫폼 활용 분석 과정 및 결과를 비교 분석하였다. 이상의 사례 및 비교 분석과정에서 나타난 도서관 빅데이터 플랫폼 활용의 한계점과 개선방안을 제시하였다.

사례분석 결과 빅데이터 활용유형분석 결과, 프로그램 기획 및 수행, 장서, 수서, 기타의 유형으로 데이터를 활용하였고 이용자 맞춤형 테마서가 및 독서진흥프로그램 진행, 장서활용도 증대, 수서및 데이터공개와 서비스들이 도출되었다. 비교분석결과, ILUS는 자관의 자료실현항분석에 특화되어 있으며, 빅데이터플랫폼은 다양한 속성(연령, 성별, 지역, 대출시기 등)에 따른 선택적 분석이 가능하여 분석시간단축과 유연한 분석이 가능하다.

도서관 빅데이터 사업과 관련하여 향후 연구 과제 및 방향을 제시하면, 첫째, 도서관뿐만 아니라 지역내 기록관, 박물관과의 정보연계 및 소장도서 구축연계로 하나의 플랫폼을 통하여 통합된 문화빅데이터 서비스 혜택을 받을 수 있는 서비스 구현을 제안한다. 둘째, 도서관 빅데이터로 구축된 도서관정보나루에서 공공도서관 문화프로그램, 평생학습프로그램, 독서진흥 프로그램 등 다양한 도서관 프로그램 콘텐츠의 통합 관리로 강사, 운영내용, 후기정보 등을 분석하고 공유할 수 있는 시스템 구현으로 지역별 특성에 맞게 적절한 프로그램을 추천해주는 도서관 운영 프로그램 관련 콘텐츠 큐레이션 서비스를 제안한다. 셋째, 도서관 업무를 공유하여 업무 수행에 도움을 줄 수 있는 매뉴얼, 연구회 보고서, 해외도서관 견학보고서, 직원역량을 강화할 수 있는 교육프로그램 자료, 도서관 통계자료, 예산 설명서, 통계 분석 툴 사용법 등 사서가 함께 공유하면 좋을 정보를 도서관 업무 데이터 플랫폼으로 구축하여 체계적이고 효율적인 업무를 수행하고, 이것을 기반으로 도서관 이용 만족도 향상과 이용자 서비스를 개선해 나갈 연구 과제가 필요하다.

아직은 도서관계 빅데이터 플랫폼이 초기단계이지만 도서관 빅데이터 분석에 대한 관심제고와 도서관 업무에 대한 다각적인 적용으로 스마트한 업무를 수행할 필요가 있다. 과학적

인 분석데이터를 근거로 이용자가 만족하는 서비스를 개발하고, 변화하는 도서관 환경에 적극적으로 대처하기 위한 도서관 빅데이터의 개발 및 확장이 기대된다.

참 고 문 헌

김서 (2013). KAIST 도서관 웹로그 분석시스템을 중심으로. 한국전문도서관협의회지, 5, 45-52.

김선태 (2013). 빅데이터의 시대: 도서관과 사서의 새로운 역할. 한국전문도서관협의회지, 5, 36-44.

김완중 (2016. 11). 공공도서관 빅데이터 분석과 의미. 한국문헌정보학회 학술발표논문집, 67-86.

김태영, 백지연, 오효정 (2018). 빅데이터 로그 기반 도서관 이용자 및 대출 현황 분석-국립세종도서관을 중심으로. 한국도서관·정보학회지, 49(2), 357-388.
<http://dx.doi.org/10.16981/kliss.49.201806.357>

심원식, 배경재, 박희진, 서유진, 안혜연, 서규원 (2015). 도서관자료로서의 빅데이터와 국립중앙도서관의 역할. 서울: 국립중앙도서관.

이수상 (2014). 도서관과 빅데이터 분석. 圖書館文化, 55(8), 14-25.

이원상 (2014). 학술정보영역의 빅데이터 플랫폼의 이해 및 구축전략. 디지털도서관, 74, 25-35.

이정미 (2013). 빅데이터 이해와 도서관 정보서비스에의 활용. 한국비블리아학회지, 24(4), 53-73.
<http://dx.doi.org/10.14699/kbiblia.2013.24.4.053>

이혜진, 김완중, 김혜선 (2015). 도서관 빅데이터 활용을 위한 분석 플랫폼 구축 및 서비스 방안-도서관 빅데이터 분석 활용 체계 구축사례를 중심으로. 디지털도서관, 79, 3-14.

조영임 (2013). 빅데이터의 이해와 주요 이슈들. 한국지역정보화학회지, 16(3), 43-65.
<https://doi.org/10.22896/karis.2013.16.3.002>

표순희 (2015. 5). 공공도서관 빅데이터 서비스의 경제적 가치. 한국비블리아 발표논문집, 7-28.

표순희, 김윤형, 김혜선, 김완중 (2015). 도서관 빅데이터 서비스 모형 개발에 관한 연구: 공공 도서관을 중심으로. 정보관리학회지, 32(2), 53-86. <https://doi.org/10.3743/KOSIM.2015.32.2.063>

한중희, 권상일, 김현숙, 서영주, 김남규 (2014). 입법정보지원연구과제보고서: 빅데이터 분석기법을 활용한 국회도서관 학위 논문 콘텐츠 강화 방안. 서울: 국회도서관.

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

Cho, Young Im (2013). Understanding big data and its main issues. Journal of Korean Association

- for Regional Information Society, 16(3), 31-52. <https://doi.org/10.22896/karis.2013.16.3.002>
- Han, Jong-Hee, Kwon, Sang-Il, Kim, Hyun-Sook, Suh, Young-Joo, & Kim, Nam-Gyu (2014). Legislative information support research project report: A proposal for thesis and dissertation contents enhancement using big data analysis in national assembly library. Seoul: National Assembly Library.
- Kim, Seo (2013). Focusing on the KAIST library web log analysis system. Korea Special Library Association(KSLA) Bulletin, 5, 45-52.
- Kim, Sun Tae (2013). Big data era: Challenging role of library as an librarian. Korea Special Library Association(KSLA) Bulletin, 5, 36-44.
- Kim, Tae-Young, Baek, Ji-Yeon, & Oh, Hyo Jung (2018). An analysis of library user and circulation status based on bigdata logs a case study of national library of Korea, Sejong. Journal of Korean Library and Information Science Society, 49(2), 357-388. <http://dx.doi.org/10.16981/kliss.49.201806.357>
- Kim, Ywan Jong (2016. 11). Analysis and meaning of big data in public libraries - Focusing on the collection recommender algorithm. Proceeding of Conference of Korean Society of Library and Information Science, 67-86.
- Lee, Hye Jin, Kim, Wan Jon, & Kim, Hyesun (2015). Big data platform implementation and service plans in libraries: Focus on the "study on big data analysis and adoption in library". Digital Libraries, 79, 3-14.
- Lee, Jeong-Mee (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), 53-73. <http://dx.doi.org/10.14699/kbiblia.2013.24.4.053>
- Lee, Su Sang. (2014). Library and big data analysis. KLA Journal, 55(8), 14-25.
- Lee, Won Sang (2014). Big data platform in academia and its implementing strategy. Digital Libraries, 74, 25-35.
- Pyo, Soon Hee (2015. 5). Economical value of big data services for public library. Proceeding of Conference of Korean Society of Biblia, 7-28.
- Pyo, Soon Hee, Kim, Yun Hyung, Kim, Hye Sun, & Kim, Wan Jong (2015). A study on the developing of big data services in public library. Journal of the Korean Society for Information Management, 32(2), 53-86. <https://doi.org/10.3743/KOSIM.2015.32.2.063>
- Shim, WonSik, Bae, Kyung-Jae, Park, Heejin, Suh, Yoo Jin, Ahn, Hye-Yeon, & Suh, Kyu-Won (2015). Big data as library contents and the role of the national library of Korea. Seoul: National Library of Korea.

[부록 1] 오류데이터 예시(최다대출 장서 순위목록/2016년)

▪ 대출장서 순위목록(0-5세)

순위	서명 (출판년도) 저자/역자 출판사	도서관 ISBN 표준 ISBN	Vol	주제분야	소장 권수	대출 빈도	
1	(똑똑한 선택을 이끄는 힘) 냇지 (2013) 이소영 옮김 서울 :예림아이	9788965673675 9788965673675			843	5	15
2	예 & 레 (2014) 메오미 지음 ; 홍연미 옮김 파주 :문학수첩 리틀북	9788959761593 9788959761593			843	4	14
3	벚꽃 흩날리는 밤 (2013) 캐서린 하프카 글 ; 그레이스 리 그림 ; 팍정아 옮김 서울 :애플비	9788926210048 9788926210048	1	808.3	4	13	
4	세계 추리소설 걸작선 2 (2012) 지나 잉골리아 개작 ; 오현주 옮김 서울 :예림아이	9788965671671 9788965671671	16	808.9	4	13	
5	노자 3 : 학문이 끝나는 곳에 도가 있다 (2011) GIMC, DPS [공] 글·그림, 백희나, 김향수 [공] 원작 한솔수북	9788953582750 9788953582750	3	808.3	3	13	
6	내가 나를 낮게 한다 (2012) 밸러리 토머스 글 ; 코키 풀 그림 ; 노은정 옮김 비룡소	9788949112312 9788949112312	229	808.3	4	13	
7	농지·산지 사업관련 법령편람 : 산지편 .2 (2014) 키즈아이콘 기획 키즈아이콘	9788964132609 9788964132609	7	813.608	2	11	
8	응답유사 (2013) GIMC, DPS [공] 글·그림 한솔수북	9788953588264 9788953588264	13	808.3	2	11	
9	Kings & Queens of England: Every Question Answered (2011) GIMC, DPS [공] 글·그림, 백희나, 김향수 [공] 원작 한솔수북	9788953582767 9788953582767	4	808.3	3	11	
10	10년 후 미래 : 세계 경제의 운명을 바꿀 12가지 트렌드 (2013) GIMC, DPS [공] 글·그림 ; 백희나 ; 김향수 [구름빵 원작] 서울 :한솔수북	9791195140053 9791195140053	16	808.3	2	11	

▪ 대출장서 순위목록(6-7세)

순위	서명 (출판년도) 저자/역자 출판사	도서관 ISBN 표준 ISBN	Vol	주제분야	소장 권수	대출 빈도
1	우리는 공산당이 싫어요 (2012) 김현수 아울북	9788950936358 9788950936358	21	711.47	3	15
2	씨앗 받는 농사 매뉴얼 (2013) 토머스, 밸러리 비룡소	9788949112459 9788949112459	227	808.3	2	13
3	벚꽃 흩날리는 밤 (2013) 캐서린 하프카 글 ; 그레이스 리 그림 ; 팍정아 옮김 서울 :애플비	9788926210048 9788926210048	1	808.3	4	12
4	한글 우수성과 한글 세계화 (2012) 김현수 아울북	9788950935252 9788950935252	20	711.47	3	11
5	내일도 행복이다 (2012) GIMC ; DPS [공] 글·그림 한솔수북	9788953584044 9788953584044	8	808.3	2	11
6	원초적 본능 feat. 미소년 (2010) 강우홍 글, 김달성 그림 서울:보림	9788943302719 9788943302719	18	388.308	2	11
7	행태주의 마케팅 (2010) 김현민 글송이	9791170181323 9791170181323	8	400	3	10
8	노래하는 사장님 (2015) 봉현주	9788966452293 9788966452293	5	331.17	2	10
9	응답유사 (2013) GIMC, DPS [공] 글·그림 한솔수북	9788953588264 9788953588264	13	808.3	2	10
10	이근성의 통기타 교실 중급편 2 (2010) 코믹컴 글 ; 네오 그림 코믹컴	9788994110059 9788994110059	6	400	2	10

▪ 대출장서 순위목록(청소년14-19세)

순위	서명 (출판년도) 저자/역자 출판사	도서관 ISBN 표준 ISBN	Vol	주제분야	소장 권수	대출 빈도
1	음로 이야기 (2010) 샌델, 마이클 파주 :김영사	9788934939603 9788934939603		340.1	7	23
2	디데이 (2007) 팀 보틀러 지음 ; 정해영 옮김 서울 :다산책방	9788992555487 9788992555487		843	3	18
3	과로 사회 (2011) 김애란 파주 :창비	9788936433871 9788936433871		813.6	7	17
4	거의 맞음 (2010) 김난도 지음 서울 :뽀앤파커스	9788965700036 9788965700036		199.1	4	17
5	전환기 중국의 정치경제 (2009) 김려형 파주 :창비	9788936456221 9788936456221		813.6	5	17
6	절대 혼자 살지 마라 (2007) 장 지글러 지음 ; 유영미 옮김 서울:갈라파고스	9788990809179 9788990809179		335.7	4	16
7	김밥 주먹밥 씹밥 (2009) 구병모 지음 파주 :창비	9788936433697 9788936433697		813.6	4	15
8	민법요해. 2: 채권총론 채권각론 (2012) 미하엘 연데 지음 ; 한미희 옮김 서울:비룡소	8949190028 9788949190020		853	5	14
9	여행사실무 .[1] (2012) 윤태호 위즈덤하우스	9788960865594 9788960865594	2	657.1	5	13
10	왕코와 백석이 (2008) 신경숙 지음 파주 :창비	9788936433673 9788936433673		813.6	4	13

[부록 2] 정정데이터 예시(최다대출 장서 순위목록/2016년)

▪ 대출장서 순위목록(0~5세)

순위	서명	출판년도	저자/역자	대출빈도
1	(Disney) 겨울왕국: movie storybook	2013	이소영 옮김	15
2	바다탐험대 옥토넷: 깊고 넓은 신비한 바다 여행	2014	메오미 지음; 홍연미 옮김	14
3	구름빵: 그림자가 사라졌어	2011	GIMC, DPS [공] 글·그림, 백희나, 김향수 [공] 원작	13
4	물란	2012	지나 잉폴리아 개작; 오현주 옮김	13
5	마녀 위니의 공룡 소동	2012	밸러리 토머스 글; 코키 폴 그림; 노은정 옮김	13

▪ 대출장서 순위목록(6~7세)

순위	서명	출판년도	저자/역자	대출빈도
1	마법천자문 .21, 하나가 되는 마음! 마음 심	2012	김현수	15
2	용감한 해적 마녀 위니	2013	토머스, 밸러리	13
3	(리틀 프린세스) 소피아	2013	캐서린 허프카 글; 그레이스 리 그림; 팍정아 옮김	12
4	호랑이 잡은 피리	2010	강무홍 글, 김달성 그림	11
5	마법천자문 .20, 내 모든 걸 너에게! 남길 유	2012	김현수	11

▪ 대출장서 순위목록(14~19세)

순위	서명	출판년도	저자/역자	대출빈도
1	정의란 무엇인가: 허버드대 20년 연속 최고의 명강의	2010	샌델, 마이클	23
2	리버보이	2007	팀 보올러 지음; 정혜영 옮김	18
3	우아한 거짓말: 김려령 장편소설	2009	김려령	17
4	아프니까 청춘이다: 인생 앞에 홀로 선 젊은 그대에게	2010	김난도 지음	17
5	두근두근 내 인생: 김애란 장편소설	2011	김애란	17