

# 대체자료 선정을 위한 이용자 참여형 도서 추천 큐레이션 플랫폼 설계\*

## Design of the Curation Platform for User-participated Book Recommendation System of Selecting on Alternative Material for the Disabled

조 현 양 (Hyun-Yang Cho)\*\*

### 목 차

- |                   |                        |
|-------------------|------------------------|
| 1. 서론             | 4. 대체자료 이용 데이터 수집 및 분석 |
| 2. 이론적 배경         | 5. 결론                  |
| 3. 이용자 분석을 위한 모델링 |                        |

### 초 록

이 연구의 목적은 이용자 선호도 및 관심분야 정보 수집을 위한 웹 기반의 프로파일링 모델을 활용하여 이용자 선호도 기반의 대체자료 추천 시스템의 개발과 정보 추천 모델의 핵심이 되는 도서자동분류엔진을 설계하고 시범 구축하는 것이다. 이용자 선호도 기반의 대체자료 선정 방법론을 개발하기 위해 국가대체자료공유시스템인 DREAM의 이용자 대체자료 이용 내역을 분석하여 KDC의 10개 주제 분야별 대체자료에 대한 이용자 선호도를 조사하였다. 이용 내역 데이터의 표제 정보를 분석하여 주제 분야별로 그리고 조사대상 기간별로 관심 키워드를 도출하였으며, 도출된 키워드를 주제 분야별로 순위화하였다. 주제별과 기간별로 순위화된 키워드는 이용자 선호도를 나타내는 것으로 출판된 도서 가운데 대체자료 제작 대상을 선정하는데 활용할 수 있다. 또한 이용자 선호도 기반의 대체자료 선정을 위하여 국립장애인도서관의 내부 데이터뿐만 아니라 타 도서관이나 다양한 유관 기관의 서비스 및 이용 내역을 반영하기 위한 연계방안을 제안하였다.

### ABSTRACT

The purpose of this study is to design and develop a alternative material recommendation system using automatic classification, based on user preference. Details of usage data by users from DREAM was analysed in order to develop the way of a method on selecting proper alternative material, and then the data by user preference were allocated under each category of 10 KDC categories. The keyword, selected from the title of users' usage data from a certain period of time, were divided into 10 subject categories and ranked by the order of frequency of appearance. Books including high frequency of the keyword in title can be selected as a preferred target for producing alternative materials. Lastly, a dynamic linkage for sharing usage data among National Library for the Disabled and other libraries is proposed to produce more proper alternative materials, based on user preference.

키워드: 대체자료, 이용자 선호도, 도서자동분류, 도서추천시스템

Alternative Material, User Preference, Automatic Classification, Book Recommendation System

\* 이 논문은 2018학년도 경기대학교 연구년 지원에 의하여 수행되었음.

\*\* 경기대학교 인문대학 문헌정보학과 교수(hycho@kyonggi.ac.kr / ISNI 0000 0004 6001 9205)

논문접수일자: 2020년 8월 3일 최초심사일자: 2020년 8월 8일 게재확정일자: 2020년 8월 18일

한국문헌정보학회지, 54(3): 41-69, 2020. <http://dx.doi.org/10.4275/KSLIS.2020.54.3.041>

## 1. 서론

### 1.1 연구의 필요성 및 목적

최근 데이터를 활용한 상업적 가치 창출 사례가 증가함에 따라 빅데이터에 대한 관심이 급격하게 증가하고 있으며, 소비자의 성향을 반영하여 개인에게 적합한 서비스를 제공하는 개인 맞춤형 서비스에 대한 관심이 증가하고 있다. 특히 개인의 성향을 반영한 독서활동이 인성 및 인격 향상에 긍정적인 영향을 준다는 연구결과가 보고됨에 따라 개인화 기반의 서비스 지원 체제가 중요하게 부각되고 있으며, 도서관에서의 이용자 맞춤형 서비스에 대한 관심도 지속적으로 증가하고 있다. 이에 따라 공공도서관을 중심으로 빅데이터 분석을 통해 이용자의 성향을 파악하여 개인에게 적합한 서비스를 제공하기 위한 다양한 접근들이 시도되고 있다.

개인 맞춤형 도서제공을 위한 다양한 접근 방법 가운데 하나로 도서관 이용자들의 개인별 정보 요구나 성향 분석을 바탕으로 한 도서 큐레이션에 대한 관심과 연구가 활발하게 진행되고 있다. 특히 멀티미디어 자료 및 기기의 확대와 디지털 자료에 익숙한 이용자들의 관심을 다시 책으로 돌리기 위하여 도서관은 이용자 모델링 및 개인화 서비스 등 다양한 측면에서 노력을 기울이고 있다.

그러나 맞춤형 서비스가 절대적으로 요구되는 장애인 이용자의 특성에도 불구하고 장애인을 위한 대체자료 제작 방법론의 부재로 인하여 대체자료의 부족으로 연계되고 있으며, 장애인들은 상대적으로 신간 도서 정보 및 전문

적인 전공 관련 정보의 획득이 어려운 것이 현실이다. 따라서 장애인도서관은 이용자의 요구를 미리 예측하여 대체자료를 제작하고 이용자 서비스를 제공하기 위한 방법론을 구축해야 하며, 이를 위해 도서관이 보유하고 있는 대출 현황과 같은 이용 데이터를 활용하여 이용자 요구를 분석하기 위한 체제 및 기술 도입이 필요하다.

이 연구의 목적은 국립장애인도서관의 이용자 데이터 및 이용 내역 데이터를 기반으로 이용자의 선호 주제 분야 및 관심 키워드를 분석하고, 이용자 관심 정보를 반영할 수 있는 도서 자동분류엔진의 설계와 장애인들이 대체자료를 선택함에 있어서 개인에게 적합하고 필요한 도서를 합리적으로 제작하고 제공할 수 있도록 이용자 선호도 기반 도서추천시스템을 제안하는 것이다.

### 1.2 연구의 내용 및 범위

본 연구를 수행하기 위한 내용 및 범위는 다음과 같다.

첫째, 개인 이용자의 성향이나 선호도 기반의 대체자료 추천 시스템을 개발하기 위하여 빅 데이터를 활용한 이용자 맞춤형 정보서비스 및 도서자동분류와 관련된 선행연구에 대한 조사를 실시하였다.

둘째, 국립장애인도서관의 현재 이용자 프로파일링 방법과 이용자 선호도 관심분야 정보 수집을 위한 웹 기반의 프로파일링 모델을 분석하였으며, 대체자료 이용 현황 데이터를 활용하여 이용자 선호도 기반 도서자동분류엔진을 설계하고, 시범 구축하였다.

셋째, 최근 6년간 장애인도서관 이용자의 대체자료 이용 내역을 분석하여 이용자 선호 주제 분야 및 분야별 관심 키워드를 도출하였다. 이를 위하여 우선 대체자료공유시스템인 DREAM의 이용내역 데이터를 이용하여 KDC 10개 주제 분야별 대체자료에 대한 이용자 선호도를 조사하였다. 그리고 내역 데이터의 표제 정보를 분석하여 주제 분야별로 관심 키워드를 도출하였으며, 도출된 키워드를 주제 분야별로 순위화하였다.

넷째, DREAM 이용 내역과는 별개로 국립장애인도서관 소장 대체자료에 대한 이용내역 데이터를 분석하여 동일한 방법으로 키워드를 순위화 하고, 2집단간 선호 키워드를 비교, 분석하였다.

마지막으로 이용자 선호도 기반의 대체자료 제작을 위한 대상 선정 방법론과 도서추천시스템을 제안하였다.

(국립장애인도서관 2007), 『지체장애인의 도서관 이용활성화를 위한 교육 및 문화프로그램 모델 개발』(국립장애인도서관 2012), 『2019 도서관장애인서비스 현황조사』(국립장애인도서관 2018), 그리고 『장애인서비스 지원 사업』(국립장애인도서관)에 이르기까지 장애인을 위한 도서관서비스의 확대 기반은 충분히 조성되었다. 특별히 제3차 도서관발전종합계획의 4대 전략 방향 가운데 하나인 ‘사회적 포용을 실천하는 도서관’은 적극적 정보복지 실현을 위하여 국가 차원에서의 정보 불평등을 완화하고 도서관 아웃리치의 지속적 확대를 추구하고 있음을 공표한 것으로 볼 수 있다.

이러한 노력에도 불구하고 2019년 기준 국내 연간 도서 출판량 대비 대체자료 제작 비율은 <표 1>에 나타난 바와 같이 12.3%(58,635 중 7,233권)에 불과하며, 특히 장애인이 접근 가능한 전자책은 2%(95,581건 중 1,961건)<sup>1)</sup> 수준으로 매우 낮은 상황이다. 또한 국민 1인당 도서관 장서 규모 2.15권에 비해 시각장애인 1인당 장서 규모 1.3권, 청각장애인 1인당 0.2권, 발달장애인 1인당 0.002권으로 극히 저조한 것이 현실이다.

공공도서관 장애유형별 장애인 등록 회원 수는 시각장애인의 경우 2016년 3,737명에서 2018년 14,168명으로, 청각장애인은 2016년 926명

## 2. 이론적 배경

### 2.1 대체자료의 서비스와 활용

1975년 UN의 장애인 권리선언으로부터 『도서관의 장애인서비스 기준 및 지침 제정 연구』

<표 1> 국내 연간 도서출판량 대비 대체자료 제작 현황

| 구 분           | 2017                | 2018                | 2019                |
|---------------|---------------------|---------------------|---------------------|
| 연간 도서출판량(종수)  | 53,795              | 56,809              | 58,635              |
| 대체자료 제작현황(건수) | <b>7,149(13.2%)</b> | <b>7,210(12.6%)</b> | <b>7,233(12.3%)</b> |
|               | 시각 6,618/청각 531     | 시각 6,451/청각 759     | 시각 6,890/청각 343     |

1) 자료 출처: 대한출판문화협회, 국립장애인도서관 보도자료.

에서 2018년 112,051명으로, 그리고 지체장애인은 4,850명에서 8,216명으로 급증한 것으로 나타났다. 장애유형별 2016년에서 2018년의 사이의 대체자료 이용자 수는 시각장애인 73,119명에서 2018년 131,471명으로, 청각장애인은 44,198명에서 85,905명으로 그리고 지체장애인은 168,775명에서 255,830명으로 급격히 증가하고 있다. 반면, 장애인 코너나 장애인자료실 등 장애인을 위한 공간이 설치되어 있는 공공도서관의 비율은 2016년 35.8%(908개관 중 325개관), 2018년 35.4%(914개관 중 324개관)<sup>2)</sup>로 소폭 감소한 것으로 나타났다.

잠재 이용자의 증가와 공공도서관에서의 장애인을 위한 장소나 시설의 확보 상태, 그리고 부족한 대체자료 문제를 해결하고 장애인에 대한 정보접근성을 향상시킬 수 있는 방안은 충분한 대체자료의 생산, 정보자원의 디지털화, 그리고 인터넷을 통해 서비스 가능하도록 하는 것이다. 특히 장애인을 위한 대체자료의 생산은 비용과 노력 등 경제적으로 막대한 비용이 투입되는 것으로 생산된 자료의 활용을 보장하기 위해서는 요구에 부합되는 도서를 대체자료로 작성하는 것이 우선되어야 한다. 따라서 출판되는 도서 가운데 특정한 기준에 따라 적합한 도서를 선정하여 대체자료로 제작하는 것이 중요하며, 이용자 선호도나 사회적 이슈 등 대체자료 제작 대상 선정을 위한 다양한 기준이 마련되어야 한다.

국립장애인도서관을 비롯하여 공공도서관은 장애인들의 정보 접근과 독서환경을 지원하고 교육기회를 제공하기 위한 다양한 서비스를 개

발, 제공하고 있다. 독서활동을 지원하기 위해 장애인 이용자와 함께 책을 읽거나 읽어 주는 낭독 프로그램을 운영하기도 하고 독서보조기를 대여해 주는 서비스를 제공하기도 한다. 대표적인 예로는 LG상남도서관이 유비쿼터스 시각장애인도서관으로 운영하고 있는 ‘책읽어주는 도서관’이 대표적이다.

국립장애인도서관은 장애인 대체자료의 공유 및 활용을 통해 장애인의 정보 접근성 향상과 지식정보 격차 해소를 위하여 국가대체자료 공유시스템(DREAM, 이하 DREAM)을 운영하고 있다. DREAM은 국립장애인도서관을 중심으로 전국의 공공도서관과 사립장애인도서관 등 45개 기관들이 보유한 대체자료 목록 및 원본을 수집하여 장애인 이용자가 통합 검색과 다운로드, 그리고 열람할 수 있는 지식정보서비스로 일부 원문을 이용할 수 있도록 지원하는 국가 대체자료공유시스템이다.

## 2.2 선행 연구

문헌정보학 분야에서 이용자의 정보 요구 충족을 위한 맞춤서비스나 적합한 도서의 추천에 대한 다양한 연구가 수행되고 있으며, 사회전반에 걸친 빅데이터의 활용과 더불어 도서관에서도 데이터를 기반으로 이용자 맞춤 문화행사 개최, 맞춤형 서비스의 개발 등 다양하게 활용되고 있다. 반면, 다량의 출판 도서 가운데 장애인 이용자들이 선호하는 도서를 우선적으로 대체자료로 제작하기 위하여 장애인의 관심 주제나 적성을 반영한 추천도서 목록을 작성할 필

2) 국립장애인도서관. 2018. 2019 도서관 장애인서비스 현황조사. p. 87, 114, 116.

요성은 인지하고 있으나 장애인 이용자의 이용 성향이나 주제 분야별 선호도를 분석하여 대체 자료 제작을 위한 도서의 선정과 관련된 연구는 수행되지 않고 있다.

윤희윤(2010)은 각국의 대체자료 개발정책을 분석하고 대체자료 개발정책의 주제, 재정 지원방식, 제작방식 및 배포, 그리고 대체자료의 서비스 등 국가도서관 주도의 대체자료 개발정책 수립을 제안하였다. 백재은(2016)은 대체자료를 이용한 독서치료의 필요성을 인식하고 장애인을 위한 독서치료용 대체자료 목록과 함께 대체자료 개발 방안을 제안하였다.

장보성 등(2012)은 우리나라 도서관 장애인 서비스의 현재 수준을 파악하기 위하여 전국 804개 공공도서관과 장애인도서관(점자도서관)을 대상으로 대체자료와 정보서비스, 문화프로그램 등의 6개 영역을 구분하여 현황을 조사·분석하였다. 분석결과, 장애인서비스의 수준과 내용에서 도서관 운영주체, 내용적 측면, 지역에 따라 다양한 편차가 있음을 발견하였다.

이정미(2013)는 도서관에서의 빅데이터 활용 방안을 모색하기 위하여 빅데이터의 활용사례와 연구 동향을 분석하고, 문헌정보학 분야에서의 빅데이터에 관심을 집중해야 하는 이유, 문제점이나 제기되는 이슈와 함께 정보서비스 분야에서의 활용 방안을 제안하였다.

권혁인 등(2014)은 도서정보와 고객 성향 기반의 도서 추천 큐레이션 서비스의 개념과 사례, 서비스 구현과정 및 비즈니스 모델을 제안하면서 도서 및 독자 SNS 연계정보의 수집, 수집된 도서DB의 통합화 및 분석, 그리고 사용자 기반 시각화 요소의 구성을 통해 고객 맞춤형 큐레이션의 효율성을 강조하였다. 김수아

등(2014)은 블로그 사이트에서 포스트의 자동 분류를 위해 분류별 문서를 수집하고, 수집된 분류별 문서의 어휘빈도와 문서빈도 그리고 분류별, 빈도 등의 다양한 어휘가중치 조합을 통하여 블로그 포스트의 특성에 적합한 가중치 방식으로 TF-CTF-IECDF를 제안하였다. 김미지 등(2018)은 TF-IDF 모델을 사용하여 단어간의 상대적 중요성에 기반하여 56,000여 개의 뉴스 기사에서 이슈가 되는 키워드를 추출하고 TF-IDF 기반의 방식과 TF 기반의 방식을 비교 분석하였으며, TF-IDF 방식을 기반으로 한 자동 추출시스템이 뉴스 기사에서의 이슈 키워드 추출에 우수한 성능을 보이는 것을 확인하였다.

안희정 등(2017)은 는 도서 서지정보, SNS 데이터, 도서 본문 텍스트 등을 분석하여 키워드로 추출하고, 도서의 주제 키워드, 배경, 등장인물과 같은 이용자로부터 획득한 정보를 기반으로 도서를 추천해 주는 이용자 맞춤형 도서 큐레이션 시스템을 제안하였다.

조현양(2017)은 개인의 성격유형에 따라 선호하는 도서에 차이가 있음을 전제로 국립어린이청소년도서관에서 제공하는 501권의 유아 및 아동도서를 대상으로 책소개 정보를 활용하여 개인별 성격유형에 적합한 도서를 합리적으로 추천할 수 있는 서평 자동분류시스템을 개발하였다. 한희준 등(2018)은 과학기술분야 연구자 개인의 관심 분야를 추적하기 위하여 개별 연구자의 연구 논문에 대한 핵심 연구 분야를 분석하여 개인별로 적합한 연구 논문을 추천할 수 있는 시스템을 개발, 제안하였다.

오선경(2017)은 자치단체의 지원 하에 점자 도서관을 중심으로 개인의 요청이나 담당자의

판단으로 시각장애인을 위한 대체자료의 제작으로 인한 대체자료 제작의 범위 제한, 중복 제작, 대체 포맷의 한정 등의 문제점을 지적하고 시각장애인과 시각장애인도서관을 대상으로 인식 조사 결과를 바탕으로 대체자료 제작 및 이용활성화를 위한 개선 방안을 제안하였다.

강성구와 임경원(2017)은 국내에서 진행된 장애인용 대체자료의 제작 현황과 기술변화를 분석하고 ePUB 포맷과 DAISY 포맷 간의 통합된 전자출판 표준 포맷의 적용에 따른 다양한 기술 변화와 이에 대응하기 위한 국가 대체자료 제작 및 서비스 역량 강화를 위한 국가장애인도서관의 개선방안을 제안하였다.

지금까지 수행된 선행연구를 바탕으로 이 연구에서는 이용자 선호도 및 관심분야 정보 수집을 위한 웹 기반의 프로파일링 모델을 활용하여 이용자 선호도 기반의 대체자료 추천 시스템의 개발과 정보 추천 모델의 핵심이 되는 도서자동분류엔진을 설계하고 시범 구축하고자 한다.

### 3. 이용자 분석을 위한 모델링

이 연구에서는 이용자 참여형 도서 추천 큐레이션 플랫폼을 설계하기 위한 방법론을 소개한다. 이를 위해 먼저 국립장애인도서관의 현재 이용자 프로파일링 기법을 조사, 분석하였으며, 국립장애인도서관의 이용자 데이터 및 대출 현황 데이터를 기반으로 이용자의 선호 주제 분야 및 관심 키워드를 분석하였다. 또한 개인화된 정보 추천 시스템을 구축하기 위하여 이용자 관심사 정보를 반영할 수 있는 정보 추

천 모델을 설계하였으며, 이를 바탕으로 모델의 핵심 기능인 도서자동분류엔진을 시범 구축하였다.

#### 3.1 주제 분야 분석을 위한 이용자 모델링

이용자 선호 주제 분야 및 관심 키워드를 분석하기 위해서는 이용자 정보를 파악하여 이를 구조화하고 조직화하는 이용자 모델링 과정이 필요하다. 따라서 국립장애인도서관의 이용자 모델링을 위해 현재 이용자 프로파일링 기법을 조사하고 분석하였다.

국립장애인도서관의 이용자 프로파일링은 대부분 웹상의 회원가입 절차를 통해 이루어지고 있으며, 국립장애인도서관의 웹 회원가입 절차를 통해 수집되는 이용자 정보는 <표 2>와 같다.

회원가입 유형은 우선 장애인 회원과 비장애인 회원으로 구분된다. 장애인과 비장애인 이용자에 대해 공통적으로 수집되는 이용자 정보는 성명, 성별, 생년월일, 회원 ID 및 비밀번호, 주소, 연락처, 이메일 주소, 회원 유형 등으로 주로 개인의 신상정보를 포함한다.

장애인 이용자의 경우 장애 등급, 장애유형, 장애인용 대체자료 서비스 이용 여부에 대한 정보가 추가로 수집되고 있으며, 대체자료인 DAISY 도서 이용을 위한 보조공학기기 등록이 회원가입 절차에서 이루어지고 있다. 장애 등급은 1등급부터 6등급, 그리고 비장애 중에서 선택하여 입력할 수 있으며, 장애 유형은 청각장애, 지체장애, 뇌병변장애, 언어장애, 시각장애, 안면장애, 신장장애, 심장장애, 간장애, 호흡기 장애, 장루/요루 장애, 뇌전증(간질) 장애, 지적장애, 자폐

〈표 2〉 국립장애인도서관 웹 회원가입을 통해 수집되는 이용자 정보

|          | 장애인 회원  | 비장애인 회원    |
|----------|---|------------|
| 이용자 신상정보 | 성명  | 성명         |
|          | 성별  | 성별         |
|          | 생년월일  | 생년월일       |
|          | 회원 ID/비밀번호                                      | 회원 ID/비밀번호 |
|          | 주소(우편번호)  | 주소(우편번호)   |
|          | 연락처   | 연락처        |
|          | 이메일 주소  | 이메일 주소     |
| 장애관련 정보  | 회원유형  | 회원유형       |
|          | 장애등급  |            |
|          | 장애유형  |            |
|          | 장애인용 대체자료 서비스 이용 여부<br>DAISY 도서 이용을 위한 보조공학기기등록 |            |

성 장애, 정신장애, 국가유공상이자, 장기요양, 기타 중에 선택이 가능하다.

회원가입 과정에서 요구되는 정보에서 확인한 바와 같이 현재 국립장애인도서관의 이용자 프로파일링은 이용자의 신상정보를 중점으로 이루어지고 있기 때문에 이용자 선호도 및 관심 분야를 확인하기 위해서는 새로운 프로파일링 기법이 필요하다. 따라서 이용자 선호도 및 관심 분야 정보를 수집하기 위한 방법으로 이 연구에서는 웹 기반의 동적 프로파일링 모델을 활용하였다.

웹 기반의 동적 프로파일링은 웹상에서 이루어지는 이용자의 다양한 활동 기록을 수집하고 이를 통해 이용자 프로파일을 구축하는 기법으로, 이용자의 추가적인 정보 기입 행위 없이 이

용자 정보를 수집할 수 있다는 장점을 가진다. 일반적인 웹 프로파일링은 〈표 3〉에 나타난 바와 같이 로그인과 같은 접속 정보, 검색 행위에 대한 로그정보인 검색 정보, 그리고 웹상에서 이루어진 다양한 이용 행위에 대한 활동 정보로 구분할 수 있다.

접속 정보는 이용자의 특성을 파악하기 위해 기본적으로 요구되며, 웹상의 행위에 대한 로그정보인 검색정보와 활동정보 수집을 통해 이용자의 이용 행태를 파악할 수 있다. 이용자 로그정보들을 활용한 이용자 선호도 및 관심분야 파악을 위하여 국립장애인도서관 웹을 통해 수집될 수 있다. 국립장애인도서관 웹을 통해 수집될 수 있는 검색정보는 키워드 검색, 표제 검색 정보가 있으며, 대체자료 다운로드 내역, 유

〈표 3〉 웹을 통해 수집 가능한 이용자 로그정보

|       |   |
|-------|---|
| 접속 정보 | 접속 일자, 접속 시간, 접속 IP, 로그인 정보 등             |
| 검색 정보 | 검색 일자, 검색 시간, 검색 키워드 정보 등                 |
| 활동 정보 | 활동 일자, 활동 시간, 열람 기록, 다운로드 기록, 연관 페이지 보기 등 |

사 문서 보기, 상세 페이지 보기, 열람 기록과 같은 활동 정보 등이다.

### 3.2 선호 기반의 정보 추천 모델 개발

이용자 선호도를 반영한 정보 추천 모델 개발을 위해 이 연구에서는 이용 내역을 기반으로 주제별로 이용자들의 선호도에 따라 도서를 자동으로 분류하고 추천하는 도서자동분류엔진을 설계하였다. 이를 바탕으로 도서 자동분류엔진을 시험적으로 구축하였으며, 엔진 구축 과정에서 프로그램의 학습 과정에 필요한 기반 데이터를 확보하기 위해 유관 데이터를 수집하고 가공·분류하는 일련의 처리 과정을 수행하였다.

#### 3.2.1 정보 추천을 위한 선호도 기반 도서자동분류엔진 개발

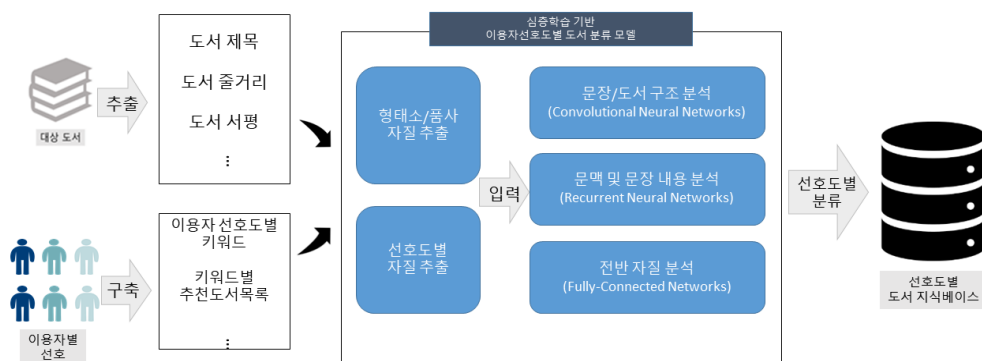
이용자의 선호에 기반하여 도서를 자동으로 분류하기 위해서는 도서를 대표할 수 있는 내적

및 외적 속성 정보를 추출할 수 있어야 한다. 예를 들어, 이용자들의 성격 유형이나 성향에 기반한 도서를 자동으로 분류하기 위해서는 우선 도서에 대한 기본 정보를 포함하고 있는 도서목록과 성격 유형별 그리고 개인의 성향을 표현할 수 있는 키워드 목록이 생성되어야 한다.

이용자 선호도를 반영하여 주제 분야나 해당 분야 내의 특정 도서를 추천하기 위한 도서 자동분류엔진의 개념은 다음의 <그림 1><sup>3)</sup>과 같다.

이 연구에서 설계된 도서자동분류 엔진은 이용자의 선호 기반에 따라 도서를 분류하기 위해 먼저 도서에 대한 속성 정보와 이용자의 선호를 표현할 수 있는 다양한 정보가 입력되어야 한다. 이를 위해 먼저 대상도서에서 도서 표제, 도서 줄거리, 도서 서평과 같은 도서 정보를 추출하고, 추출된 정보에서 형태소 및 품사적 자질을 추출한다.

다음으로 활용정도나 이용자의 관심 정도에 의해 분류된 주제 분야별 키워드나 주제분야별 추천도서 목록에서 선호와 관련된 자질을 추출한다.



<그림 1> 도서정보 추천을 위한 이용자 선호 기반 도서 자동분류 엔진

3) 국립장애인도서관. 2017. 『빅데이터 분석을 통한 대체자료 서비스 전력 수립 연구』에서 연구자가 제안한 내용과 동일함.



추출된 도서 및 선호 유형별 자질들은 선호 기반 도서 분류 모델에 입력되어 CNNs(Convolutional Neural Networks), RNNs(Recurrent Neural Networks), FC Networks(Fully-connected Networks)와 같은 딥러닝 기법을 통해 문장 및 도서 구조, 문맥 및 문장 내용, 전반 자질에 대한 분석 과정을 거치게 된다. 분석을 통해 선호도에 따라 도서가 분류되면 해당 정보는 선호 유형별 도서 지식베이스로 축적된다. 출판된 도서의 속성 정보와 이용 내역 데이터에 기반하여 장애인 이용자들이 선호하는 도서의 표제로부터 키워드를 추출, 비교하여 선호도에 따른 대체자료의 제작을 위한 대상 자료를 선정하는 것이다.

### 3.2.2 이용자 선호도 기반 대체자료 선정 방법론

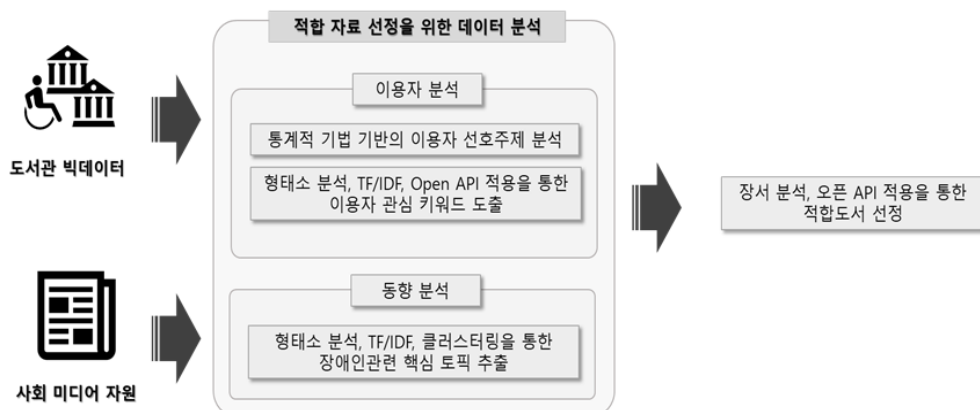
출판된 모든 문헌을 대체자료로 제작하는 것은 현실적으로 불가능하기 때문에 대체자료 대상으로 선정하기 위한 적합한 기준이 필요하다. 대체자료 선정을 위한 기준으로는 이용자 선호

도와 이용정도가 기초자료로써 활용이 우선적으로 가능하지만 그 외에도 사회적 이슈가 되는 도서, 권장도서, 추천도서, 도서 대출빈도, 베스트셀러 등 다양한 기준의 적용이 가능하다. 이 연구에서는 대체자료 제작을 위한 기준으로 대체자료 이용 기반 이용자 선호도를 반영하였으며, 이에 따라 대체자료 제작 대상 도서를 선정하기 위한 방법론을 제안하였다.

이용자 선호도 분석 결과를 반영한 대체자료 제작 대상 도서 목록을 구성하기 위한 대체자료 제작 대상 선정을 위한 방법론에 대한 개요는 다음의 <그림 2>와 같다.

제안된 방법론은 도서관 빅데이터 및 사회 미디어 자원을 활용하여 이용자 선호도와 장애인 관련 분야의 동향을 분석하여 핵심 토픽 및 키워드를 추출하고, 해당 키워드를 Open API에 활용하여 대체자료 제작에 적합한 도서를 선정한다.

먼저 이용자 선호도 분석을 위해 통계 분석과 형태소 분석, TF-IDF 가중치 계산, Open API 기술 등을 활용하고, 동향 분석을 위해 형



<그림 2> 이용자 선호도 기반의 대체자료 선정 방법론 개요

태소 분석과 TF-IDF 가중치 계산, 클러스터링 기법을 활용한다. 분석을 통해 이용자의 선호 주제 및 관심 키워드와 장애인 관련 분야 핵심 토픽이 추출되면 이를 통해 도서관 및 대형 서점, 포털 사이트 등의 Open API을 활용하여 키워드 및 토픽을 검색하고 관련된 도서 정보를 수집한다. 수집 과정에서 도서관 다대출도서 정보나 대형서점의 베스트셀러 도서 정보들을 활용하면 범용 분야의 이용자들이 가장 많이 이용하는 도서 정보 수집이 가능하다. 이와 같은 방법으로 대체자료 제작 대상 리스트가 구성되면, 다시 국립장애인도서관의 소장 대체자료 정보를 분석하여 이미 제작되어 서비스 중에 있는 대체자료를 찾아 중복 장서를 제거하여 최종적으로 대체자료 제작 대상 리스트를 구성하는 단계를 거치게 된다.

#### 4. 대체자료 이용 데이터 수집 및 분석

장애인 이용자의 정보 요구에 부응하는 대체자료 개발을 위해 우선 국립장애인도서관 이용자의 대출 이력 데이터를 수집하였으며, 이를 바탕으로 이용자 선호 주제 분야 및 키워드를 분석을 실시하였다. 이용자 선호도를 반영하여 대체자료 제작 대상 도서를 선정하기 위한 방법론을 제안하였다.

##### 4.1 DREAM 이용내역 수집

이 연구에서 이용자 선호도 분석을 위해 사용한 데이터는 DREAM의 이용 내역 데이터로,

해당 데이터에는 국립장애인도서관을 비롯하여 DREAM에 참여하고 있는 46개 기관의 대체자료 이용 내역이 구축되어 있다. DREAM은 국립장애인도서관을 중심으로 전국의 공공도서관, 사립장애인도서관들이 보유한 대체자료 목록 및 원본을 수집하여 장애인 이용자가 한번에 통합 검색, 다운로드, 열람할 수 있는 지식정보서비스이다.

수집 대상은 2013.03.07일부터 2019.09.30일까지 약 6년 6개월 동안의 국립장애인도서관 대체자료 다운로드 이력 데이터이다. 개별 데이터는 이용 정보(다운로드 날짜, 사용자 ID, 다운로드 기기), 대체자료 정보(보유 도서관명, 표제, 저자, 출판사, 출판년, 대체자료 제작일, 분류번호 등) 등을 포함하고 있다. 데이터의 분석을 위하여 해당 기간을 3개 구간으로 구분하였으며, 이 연구에서는 편의상 2013년 3월부터 2015년 12월까지를 'A' 기간, 2016년 1월부터 2017년 10월까지를 'B', 그리고 마지막으로 2017년 11월부터 2019년 9월까지를 'C' 구간으로 명명하였다.

다음의 <표 4>는 데이터 수집 대상 기간별 DREAM 이용 내역 데이터 건수를 나타내고 있다. 기간별 데이터 이용 내역을 살펴보면, 'A' 기간에 이용된 데이터는 108,036건, 'B' 기간 211,788건, 그리고 'C' 기간 249,121건 총 568,945에 이른다. 전체 이용 내역을 국립중앙도서관과 연계기관으로 구분하면 국립중앙도서관의 데이터 이용 건수가 332,219건으로 전체의 58.4%, 그리고 연계기관을 통한 이용 내역이 138,886건으로 24.4%, 그리고 도서관 정보를 포함하지 않은 데이터가 97,840건으로 17.2%이다. 여기에서 '연계기관'의 이용 내역은 DREAM에 참

〈표 4〉 대체자료 소장도서관별 DREAM 이용 내역 데이터 건수

| 기간                              | 대체자료소장도서관 | 데이터 건수  | 데이터 건수(합계) |
|---------------------------------|-----------|---------|------------|
| 전체(A+B+C)                       | 국립중앙도서관   | 332,219 | 568,945    |
|                                 | 연계기관      | 138,886 |            |
|                                 | 도서관 정보 없음 | 97,840  |            |
| 2013.03.07.<br>~ 2015.12.31.(A) | 국립중앙도서관   | 55,982  | 108,036    |
|                                 | 연계기관      | 2,420   |            |
|                                 | 도서관 정보 없음 | 49,634  |            |
| 2016.01.01.<br>~ 2017.10.31.(B) | 국립중앙도서관   | 106,683 | 211,788    |
|                                 | 연계기관      | 56,899  |            |
|                                 | 도서관 정보 없음 | 48,206  |            |
| 2017.11.07.<br>~ 2019.09.30.(C) | 국립중앙도서관   | 169,554 | 249,121    |
|                                 | 연계기관      | 79,567  |            |
|                                 | 도서관 정보 없음 | 0       |            |

여하고 있는 도서관 및 복지관 가운데 국립중앙도서관을 제외한 전체 46개 참여기관 가운데 실제 자료를 공유하는 30여개 기관을 통하여 이용된 건수를 의미한다.

기간 'C'에 해당하는 분석대상 데이터는 이용정보 내에 도서 정보를 포함하고 있으며, 기간 'A'와 'B' 기간에 포함된 분석대상 데이터는 도서 정보 데이터와 이용 정보 데이터가 분리되어 존재한다. 따라서 이용 정보 데이터에 존재하는 대체자료명을 도서정보 데이터에 검색하여 해당 대체자료를 구축한 소장도서관 정보를 확인하였다. 이 과정에서 도서 정보 데이터에 대체자료명을 포함하지 않은 관계로 검색이 되지 않은 대체자료가 A기간에 49,634건, B기간에 48,206건 존재하였으며, 이 데이터는 분석대상에서 제외하였다. 이와 유사한 이유로 수집된 데이터 가운데 분석이 불가능한 데이터를 제외하기 위한 데이터 전처리 과정이 필요하다. 데이터 전처리 과정은 우선 '표제'를 대상으

로 형태소 분석을 수행하고, 명사에 해당하는 어휘를 키워드로 추출하였다. 도서명에 명사를 포함하고 있지 않은 관계로 키워드가 추출되지 않은 데이터는 제거하였다. 형태소 분석은 한국어 정보처리를 위한 파이썬 패키지 KoNLPy의 API 중에서 'Kkma'<sup>4)</sup>를 활용하였다. 다음으로 KDC로 주어진 분류 번호를 최상위의 유(類)개념으로 10개의 주제 분야로 변환하였다. 예를 들어 688.0912는 최상위 개념인 예술로 변환하여 사용하였다. 다만 분류번호가 없는 데이터는 분석 대상에서 제거하는 데이터 전처리 과정을 거쳤다.

2013년 3월부터 2019년 9월까지 대체자료 다운로드 건수는 568,945 건이며, 이 가운데 전처리 과정을 거쳐 대체자료명이나 분류번호를 포함하고 있지 않은 98,511건이 제거되었다. 이용 내역 분석을 위하여 대체자료 다운로드 이력 분석을 위하여 사용된 데이터는 총 470,434건으로 그 내용은 다음의 〈표 5〉와 같다.

4) 한국어 정보처리를 위한 파이썬 패키지 KoNLPy(<https://konlpy-ko.readthedocs.io/ko/v0.4.3/#api>)의 API 중에서 'Kkma'를 활용하였음.

〈표 5〉 대체자료 다운로드 이력 가공/정제 후 데이터 수

| 다운로드 기간                      | 가공 후 데이터 수     | 제거 데이터           |
|------------------------------|----------------|------------------|
| 2013.03.07. - 2015.12.31.(A) | 49,123         | (-58,913)        |
| 2016.01.01. - 2017.10.31.(B) | 176,906        | (-34,882)        |
| 2017.11.07. - 2019.09.30.(C) | 244,405        | (-4,716)         |
| 합계                           | <b>470,434</b> | <b>(-98,511)</b> |

## 4.2 DREAM 대체자료 이용 내역 분석

### 4.2.1 이용자 선호 주제 분야 및 대체자료 유형 분석

이용자 선호 주제 분석은 DREAM을 통한 대체자료 이용 내역 데이터에서 한국십진분류법의 10개 주류인 '총류', '철학', '종교', '사회과학', '자연과학', '기술과학', '예술', '언어', '문학', '역사'로 구분하여 주제 분야별로 해당하는 대체자료의 이용 건수를 단순 집계하는 통계분석 방법을 통해 수행하였다.

다음으로 이용자 관심 키워드 분석은 DREAM 이용 내역 데이터를 주제에 따라 구분하고, 각 주제에 대한 키워드를 추출하여 각 키워드별 중

요도를 계산하는 방법으로 수행되었다. 먼저, 키워드 추출을 위해 이용 내역 데이터의 대체자료의 표제를 대상으로 형태소 분석을 실시하였으며, 품사 정보가 명사인 단어들을 해당 대체자료의 키워드로 추출하였다. 〈표 6〉은 DREAM 이용 내역 데이터에 대한 형태소 분석 결과이다.

그러나 문학 분야의 경우 대체자료의 표제에 이용자의 관심사가 반영된 명사가 포함되지 않아 형태소 분석을 통한 관심 키워드 도출이 어렵다. 이에 따라 Open API 검색을 통해 문학 대체자료의 카테고리 정보를 수집하였다. 문학 주제 분야의 키워드 도출을 위한 Open API 검색은 포털사이트 Naver의 'Naver 개발자 센터<sup>5)</sup>를 통해 제공되는 검색 API를 적용하여 수

〈표 6〉 형태소 분석 결과 예시

|      |           |   |
|------|-----------|---|
| 예시 1 | 주제        | 사회과학  |
|      | 표제명       | 신경영조직론  |
|      | 형태소 분석 결과 | (‘신’, ‘NNG’), (‘경영’, ‘NNG’), (‘조직론’, ‘NNG’)   |
|      | 추출된 키워드   | 신, 경영, 조직론  |
| 예시 2 | 주제        | 기술과학  |
|      | 표제명       | 임상 한의사를 위한 요통치료   |
|      | 형태소 분석 결과 | (‘임상’, ‘NNG’), (‘한의사’, ‘NNG’), (‘를’, ‘JKO’), (‘위하’, ‘VV’), (‘ㄴ’, ‘ETD’), (‘요통’, ‘NNG’), (‘치료’, ‘NNG’) |
|      | 추출된 키워드   | 임상, 한의사, 요통, 치료   |

5) <https://developers.naver.com/docs/search/book/>

행하였으며, 형태소 분석을 통해 추출된 키워드는 추출된 각 키워드에 중요도를 부여하기 위해 TF-IDF(Term Frequency - Inverse Document Frequency) 가중치를 계산하였다(〈표 7〉 참조).

(1) 주제 분야별 대체자료 이용 추세 변화  
3 구간으로 구분된 기간별로 10개 주제 가운데 어떤 주제 분야가 많이 활용되었는가를 분석하였다. 기간에 따른 주제별 대체자료 이용 추세 변화를 살펴보면, 'A' 기간의 경우 총류 → 사회과학 → 문학 → 기술과학 → 예술 → 철학 → 역사 → 언어(어학) → 종교 → 자연과학(순수과학)의 순으로, 총류에 대한 이용율이 가장 높게, 그리고 사회과학과 문학이 그 뒤를 이은 것으로 나타났다. 반면 언어, 종교, 그리고 자연과학은 이용율이 낮게 나타났다. 'B' 기간에서는 문학 → 사회과학 → 총류 → 기술과학 → 철학 → 종교 → 역사 → 언어(어학) → 예술 → 자연과학(순수과학)의 순으로 나타나 'A' 기간의 이용 추세와 비교하여 이용자들의 관심의 변화와 함께 이용율의 변화가 일어나고 있음을 보여준다. 특히 문학에 대한 이용율이 가장 높게, 그리고 사회과학과 총류가 그 뒤를 이었다. 마지막으로 'C' 기간에서의 이용율 순위는 문학 → 사회과학 → 기술과학 → 철학 → 종교 → 예술 → 역사 → 총류 → 예술 → 자연과학(순수과학)(2017.01.01. - 2019.09.30.)으로 나타나 'A' 기간과 비교하

여 총류에 대한 이용율은 높은 감소율을 보인 반면에 기술과학에 대한 이용율은 증가한 것으로 나타났다.

다음의 〈표 8〉은 기간에 따른 주제별 대체자료 이용 추세 변화를 보여 주고 있다.

〈표 8〉에 나타나 바와 같이 총류를 제외한 모든 주제 분야에서 대체자료의 다운로드 건수는 급격하게 증가하고 있으며, 주제별로 구성 비율은 차이를 보이고 있다. 다운로드 건수에 기반하여 주제별 구성 비율을 살펴보면, 특히 문학 자료의 이용률은 A기간에서는 전체의 8.64%, B기간 29.57%, C기간 45.11%로 기간별로 점차 증가하였으며, C기간에서는 45% 이상을 차지하여 이용자들의 관심 증가와 함께 이용율이 크게 늘어나는 것을 확인할 수 있다. 반면 총류는 A기간에서는 전체의 54.73%, B기간 12.86%, C기간 3.02%로 기간별로 점차 감소하여 C기간에서는 전체 이용률의 3%에 머무르고 있다. 또한 철학과 기술과학의 경우 이용율이 확대되고 있으나 자연과학(순수과학) 자료는 모든 기간에서 가장 낮은 이용률을 보이고 있다. 이러한 이용율의 변화는 국립장애인도서관의 주제별 대체자료 보유 현황(비율)과 비교하여 분석할 필요가 있다.

(2) 대체자료 유형별 이용 추세 변화  
〈표 9〉는 조사대상 기간에 따른 대체자료 유

〈표 7〉 네이버 책 카테고리 예시

|    |        |  |
|----|--------|--|
| 소설 | 나라별 소설 | 한국소설, 영미소설, 일본소설, 중국소설, 러시아소설, 프랑스소설, 독일소설, 기타나라소설 |
|    | 고전/문학  | 한국고전소설, 세계문학, 세계고전                                 |
|    | 장르소설   | SF/판타지, 추리, 전쟁/역사, 로맨스, 무협, 공포/스릴러                 |
|    | 테마소설   | 인터넷소설, 영화/드라마소설, 가족/성장소설, 어른을 위한 동화, 라이트 노벨        |

〈표 8〉 주제별 대체자료 다운로드 이력 정제 후 데이터 수

| 다운로드 기간<br>분류(KDC) | A기간<br>(2013.03.07. - 2015.12.31.) | B기간<br>(2016.01.01. - 2017.10.31.) | C기간<br>(2017.11.07. - 2019.09.30.) |
|--------------------|------------------------------------|------------------------------------|------------------------------------|
| 총류                 | <b>26,887(54.73%)</b>              | <b>22,751(12.86%)</b>              | 7,388(3.02%)                       |
| 철학                 | 1,741(3.54%)                       | 12,690(7.17%)                      | 18,373(7.52%)                      |
| 종교(총교)             | 1,022(2.08%)                       | 9,519(5.38%)                       | 11,089(4.54%)                      |
| 사회과학               | <b>7,031(14.31%)</b>               | <b>34,225(19.35%)</b>              | <b>44,875(18.36%)</b>              |
| 자연과학(순수과학)         | 566(1.15%)                         | 3,793(2.14%)                       | 3,143(1.29%)                       |
| 기술과학               | 2,537(5.16%)                       | 16,345(9.24%)                      | <b>22,589(9.24%)</b>               |
| 예술                 | 2,175(4.43%)                       | 7,493(4.24%)                       | 5,199(2.13%)                       |
| 언어(어학)             | 1,331(2.71%)                       | 8,328(4.71%)                       | 10,877(4.45%)                      |
| 문학                 | <b>4,242(8.64%)</b>                | <b>52,312(29.57%)</b>              | <b>110,248(45.11%)</b>             |
| 역사                 | 1,591(3.24%)                       | 9,450(5.34%)                       | 10,624(4.35%)                      |
| <b>합계</b>          | <b>49,123</b>                      | <b>176,906</b>                     | <b>244,405</b>                     |

〈표 9〉 유형별 대체자료 다운로드 이력 가공/정제 후 데이터 수

| 다운로드 기간<br>대체자료 유형 | A기간<br>(2013.03.07. - 2015.12.31.) | B기간<br>(2016.01.01. - 2017.10.31.) | C기간<br>(2017.11.07. - 2019.09.30.) |                |
|--------------------|------------------------------------|------------------------------------|------------------------------------|----------------|
| 페이지자료              | 전체                                 | 171                                | 490                                | -              |
|                    | 텍스트                                | 3,884                              | <b>139,985</b>                     | <b>200,585</b> |
|                    | 오디오                                | 5,403                              | 2,748                              | 727            |
|                    | 스크린리더                              | 4,941                              | -                                  | -              |
| 점자자료               | 전자점자도서                             | -                                  | 311                                | 1569           |
|                    | 전자점자악보                             | <b>11,410</b>                      | 8,089                              | 800            |
|                    | 전자점자자료                             | 3,497                              | 2,498                              | -              |
| MP3녹음도서            | 44                                 | 238                                | -                                  |                |
| 디지털음성자료            | -                                  | -                                  | 68                                 |                |
| 보이스브레일자료           | 442                                | 207                                | 130                                |                |
| 수화영상도서자료           | <b>14,489</b>                      | <b>18,303</b>                      | -                                  |                |
| 시각원문자료             | 2,891                              | 2,519                              | 23,136                             |                |
| 자막영상자료             | 535                                | 298                                | -                                  |                |
| 화면해설영상자료           | 722                                | 424                                | 556                                |                |
| 휴먼음성자료             | -                                  | 780                                | 3,773                              |                |
| 전자책                | -                                  | -                                  | 13,061                             |                |
| 기타                 | 694                                | 16                                 | -                                  |                |
| <b>합계</b>          | <b>49,123</b>                      | <b>176,906</b>                     | <b>244,405</b>                     |                |

형별 이용 추세 변화를 보여 주고 있다. 대체자료 유형은 페이지자료, 점자 자료, MP3 녹음도서, 디지털음성자료, 보이스브레일자료, 수화영

상도서자료, 시각원문자료, 자막영상자료 등의 12가지로 구분하였다.

대체자료 유형별 다운로드 이력 데이터 건수

를 살펴보면, 디지털자료가 증가율과 전체 건수에서 가장 큰 변화를 보이고 있다. 특히 텍스트 디지털자료의 경우 'A' 기간에 는 전체 이용건수 49,123 가운데 3,884건으로 7.9%였으나, 'B' 기간 139,985건으로 전체의 79.1%, 그리고 'C' 기간 200,585건으로 전체의 82.1%에 해당하여 선호하는 대체자료 유형인 것으로 나타났다.

유형별 대체자료 이용 건수는 'A' 기간 49,123건, 'B' 기간 176,906건, 그리고 'C' 기간 244,405건으로 대폭 증가하였으며, 특히 유형별로 보면 시각원문자료와 휴먼음성자료, 전자점자도서, 그리고 전자책의 경우 이용율이 큰 폭으로 증가하고 있다. 그러나 전자점자자료와 보이스브레일 자료, MP3녹음도서, 그리고 자막영상자료는 다운로드 건수와 이용율이 소폭 감소하고 있는 것으로 나타났다.

#### 4.2.2 주제별 관심 키워드 분석

##### (1) 기간별 대체자료 키워드 분석

3개 구간으로 구분된 기간 동안 이용된 대체자료의 표제를 이용하여 키워드 분석을 실시하

였다. 'A' 기간 동안 이용내역 데이터에서 등장빈도 10회 이상인 키워드는 총 3,026개로 나타났다. 'A' 기간에서 가장 많이 이용된 주제 분야는 총류로 나타났으며, 전자점자악보가 가장 많이 이용된 대체자료 유형이었다.

'A' 기간의 대체자료 이용내역에서 키워드를 추출하고 등장 빈도를 순위별로 확인한 결과는 <표 10>과 같다. 키워드별 출현빈도 수 분석결과 '피아노(PIANO)', '작품집', 'MAZURKEN', 'MINOR', 'MAJOR', 'VIOLIN', '클라리넷' 등과 같이 악기명이나 악보에서 사용될만한 키워드의 등장빈도가 높은 것으로 나타났는데, 이 기간 동안에 가장 많이 활용된 대체자료 유형인 전자점자악보와 관련이 있는 것으로 보인다. 또한 두 번째로 많이 이용된 주제 분야인 사회과학과 연관성이 있는 '복지', '교육', '사회', '세상', '연구', '이해' 등의 키워드의 등장 빈도도 높게 나타났다.

'B' 기간(2016.01.01. - 2017.10.31.) 동안 이용내역 데이터에서 추출된 등장 빈도 10회 이상 키워드는 총 7,096개로 나타났다. 이 기간

<표 10> 'A' 기간 동안 이용 대체자료의 키워드 빈도

| A (2013.03.07. - 2015.12.31.) |        |      |    |          |     |
|-------------------------------|--------|------|----|----------|-----|
| 순위                            | 키워드    | 빈도   | 순위 | 키워드      | 빈도  |
| 1                             | 피아노    | 2339 | 11 | 한국       | 838 |
| 2                             | PIANO  | 1946 | 12 | 우리       | 724 |
| 3                             | 작품집    | 1636 | 13 | 성악       | 694 |
| 4                             | 사회     | 1583 | 14 | 연구       | 686 |
| 5                             | 복지     | 1411 | 15 | 이해       | 657 |
| 6                             | 교육     | 1349 | 16 | CONCERTO | 647 |
| 7                             | 이야기    | 1302 | 17 | 이론       | 620 |
| 8                             | MAJOR  | 1017 | 18 | 장애       | 609 |
| 9                             | 세상     | 938  | 19 | SOLO     | 598 |
| 10                            | SONATA | 863  | 20 | BWV      | 594 |

동안 가장 많이 이용된 주제 분야는 문학이었으며, 텍스트데이지 자료가 가장 많이 이용된 대체자료 유형이었다.

〈표 11〉은 'B' 기간의 이용내역에서 키워드를 추출하고 등장 빈도를 1위에서 20위까지의 순위대로 나열한 결과이다. 확인 결과 '소설', '장편', '무협', '이야기' 등과 같이 문학 도서에 확인할 수 있는 키워드의 등장빈도가 높은 것으로 나타났다. 또한 두 번째로 많이 이용된 주제 분야인 사회과학과 연관성이 있는 '사회', '복지', '한국', '교육', '아이', '특수' 등의 키워드의 등장 빈도도 높게 나타나고 있다. 이 기간 동안 'PIANO', '피아노', '작품집'과 같이 총류로 분류된 전자점자악보 자료와 연관성 있는 단어도 빈도가 높은 것으로 나타났지만 A기간에서 빈도 수 1위를 차지한 '피아노'의 경우 이 기간에는 22위로 순위가 부여되어 기간 간 순위를 비교해 볼 때 상당히 하락한 것으로 보인다. 이용 건수 자체도 2,339건에서 1,442건으로 감소되었다.

'C'(2017.01.01. - 2019.09.30.)기간 이용내역

데이터에서 추출된 등장 빈도 10회 이상 키워드는 총 7,616개로 나타났다. C기간에서 가장 많이 이용된 주제 분야는 B기간과 동일하게 문학이었으며, 텍스트데이지자료가 가장 많이 이용된 대체자료 유형으로 조사되었다.

〈표 12〉는 'C' 기간의 이용내역에서 키워드를 추출하고 등장 빈도를 상위 1위에서 20위까지로 나열한 결과이다. 분석결과 키워드 등장 빈도는 문학과 연관성 있는 '소설', '장편', '완결', '판타지' 등이 가장 높은 것으로 나타났으며 '교육', '사회', '치료', '복지' 등과 같이 사회과학과 관련이 있는 키워드의 등장빈도도 높게 나타났다. 특히 '기출', '개론' 등과 같이 수험서와 관련된 키워드도 상위로 등장하는 것이 하나의 현상으로 볼 수 있다. 'A', 'B' 기간과 다르게 '총통', '통증', '기술', '의학' 등의 키워드가 상위 100에 나타났는데, 'C' 기간에서 기술과학 분야 자료 이용률이 높아진 것과 건강에 대한 사회적 관심 증가와 관련이 있을 것으로 예측이 가능하다.

〈표 11〉 'B' 기간 동안 이용 대체자료의 키워드 빈도

| B (2016.01.01. - 2017.10.31.) |     |      |    |     |      |
|-------------------------------|-----|------|----|-----|------|
| 순위                            | 키워드 | 빈도   | 순위 | 키워드 | 빈도   |
| 1                             | 소설  | 8104 | 11 | 한국  | 2347 |
| 2                             | 장편  | 5468 | 12 | 이해  | 2058 |
| 3                             | 완결  | 3799 | 13 | 우리  | 1928 |
| 4                             | 판타지 | 3677 | 14 | 기출  | 1922 |
| 5                             | 사회  | 3615 | 15 | 역사  | 1918 |
| 6                             | 이야기 | 3382 | 16 | 치료  | 1898 |
| 7                             | 무협  | 3169 | 17 | 사랑  | 1776 |
| 8                             | 교육  | 3016 | 18 | 세상  | 1717 |
| 9                             | 복지  | 2928 | 19 | 세계  | 1549 |
| 10                            | 영어  | 2352 | 20 | 아이  | 1489 |



〈표 12〉 'C' 기간 동안 이용 대체자료의 키워드 빈도

| C (2017.01.01. - 2019.09.30.) |     |       |    |     |      |
|-------------------------------|-----|-------|----|-----|------|
| 순위                            | 키워드 | 빈도    | 순위 | 키워드 | 빈도   |
| 1                             | 소설  | 13388 | 11 | 복지  | 2524 |
| 2                             | 장편  | 9140  | 12 | 한국  | 2368 |
| 3                             | 완결  | 7905  | 13 | 사랑  | 2293 |
| 4                             | 판타지 | 7153  | 14 | 사람  | 2148 |
| 5                             | 무협  | 4809  | 15 | 현대  | 2090 |
| 6                             | 영어  | 3751  | 16 | 이해  | 2024 |
| 7                             | 교육  | 3587  | 17 | 기출  | 1939 |
| 8                             | 사회  | 3256  | 18 | 우리  | 1880 |
| 9                             | 이야기 | 2910  | 19 | 상답  | 1800 |
| 10                            | 치료  | 2573  | 20 | 역사  | 1702 |

(2) 기간에 따른 대체자료 주제별 키워드 변화 분석

기간 변화에 따른 이용 대체자료 10개 주제 분야 가운데 총류(000)와 철학(100) 관련 대체 자료의 키워드 변화 정도를 분석하였다. 〈표 13〉은 기간별로 유형별 대체자료 이용 건수와 주제분야별 이용율의 변화정도를 분석한 결과로, 자료의 표제에 기반한 관심 키워드의 출현 빈도에 따른 순위를 보여 주고 있다. 분석결과, 주제 분야별 키워드 순위 변화는 기간별로 유형별 대체자료 이용 건수와 주제분야별 이용율의 변화정도와 유사한 경향을 보이고 있다. 예를 들어 'A' 기간에는 주제 분야는 총류, 대체 자료 유형은 전자점자악보가 가장 높은 이용율을 보였으며, 이 기간에는 '피아노'나 '작품집' 등이 상위에 랭크되었으나 사회과학이나 문학이 높은 이용율을 보인 'B'와 'C' 기간에는 총류에서도 '독서'나 '프로그래밍'이 상위를 차지하게 된 것이다.

총류에서 나타나는 이용자 관심사를 나타내는 키워드는 'A' 기간의 '악보', '악기' 중심에서

점차 '프로그래밍', '데이터', 'IT'와 같은 컴퓨터 프로그래밍 쪽으로 이동하고 있다. 특히 '프로그래밍'이나 '독서'와 같은 키워드는 'A' 기간에는 상위 20위 이내에 나타나지 않았던 것과 비교하여 'B'와 'C' 기간에는 1위와 3위에 오를 만큼 이용율과 관심도가 급증한 것으로 분석되었다. 그 외에도 '정보', '도서관', '인문학' 등이 'B'와 'C' 기간에 총류 분야의 주요 키워드로 등장하고 있다.

철학에서는 이용자 관심사의 큰 변화가 나타나지 않고 있다. '심리', '사주', '상답' 등과 같은 키워드가 'A', 'B', 'C' 기간에서 모두 상위권을 차지하고 있는데 그 이유는 철학이라는 학문의 성격 자체가 급격한 관심의 변화를 불러오지 않을 뿐만 아니라 그 이용 대상이 한정되어 있기 때문인 것으로 보인다.

다음으로 구간으로 구분된 3개 기간의 변화에 따른 이용 대체자료 10개 주제 분야 가운데 종교(200)와 사회과학(300) 분야 대체자료의 키워드 변화 정도를 분석하였다. 〈표 14〉는 기간별로 대체자료 이용 건수와 분야별 이용율의

〈표 13〉 기간별 총류와 철학 분야의 대체자료 키워드 분석결과

|    | 총류            |             |             | 철학       |           |          |
|----|---------------|-------------|-------------|----------|-----------|----------|
|    | A             | B           | C           | A        | B         | C        |
| 1  | PIANO(1827)   | 프로그래밍(1021) | 프로그래밍(1085) | 심리학(165) | 심리학(1049) | 명리(950)  |
| 2  | 피아노(1813)     | PIANO(831)  | 파이(635)     | 사주(125)  | 사주(992)   | 사주(910)  |
| 3  | 작품집(1612)     | 독서(634)     | 독서(410)     | 심리(124)  | 심리(577)   | 심리학(907) |
| 4  | MAJOR(898)    | 정보(472)     | 데이터(369)    | 마음(88)   | 상담(530)   | 심리(863)  |
| 5  | SONATA(789)   | 작품집(471)    | 입문(331)     | 행복(69)   | 명리(511)   | 상담(775)  |
| 6  | BWV(583)      | 피아노(441)    | 안드로이드(294)  | 인생(64)   | 철학(499)   | 인생(619)  |
| 7  | CONCERTO(565) | 이야기(411)    | IT(247)     | 놀이(63)   | 청소년(458)  | 철학(584)  |
| 8  | SOLO(561)     | IT(395)     | 리닝(238)     | 상담(56)   | 한글(443)   | 행복(478)  |
| 9  | 사회(560)       | 평설(389)     | 배우(219)     | 사랑(55)   | 대운(435)   | 사람(453)  |
| 10 | 이야기(555)      | MAJOR(385)  | 정보(203)     | 가지(55)   | 새내기(412)  | 마음(432)  |
| 11 | 세상(547)       | SONATA(372) | 도서관(194)    | 분석(53)   | 명사(411)   | 치료(375)  |
| 12 | 복지(533)       | 파이(369)     | JS(179)     | 인간(52)   | 늦깎이(411)  | 가지(369)  |
| 13 | SONATAS(525)  | 지식(325)     | HTML(173)   | 습관(50)   | 학도(411)   | 인간(317)  |
| 14 | 교육(520)       | 아이(301)     | 활용(169)     | 관계(47)   | 만세력(411)  | 사랑(307)  |
| 15 | MINOR(457)    | 도서관(283)    | 인문학(167)    | 절대(42)   | 습관(408)   | 생각(302)  |

〈표 14〉 기간별 종교 및 사회과학 분야의 대체자료 키워드 분석결과

|    | 종교       |          |          | 사회과학     |           |           |
|----|----------|----------|----------|----------|-----------|-----------|
|    | A        | B        | C        | A        | B         | C         |
| 1  | IDEA(64) | 하나님(447) | 기독교(525) | 사회(943)  | 사회(3084)  | 교육(3167)  |
| 2  | MAIN(64) | 신학(430)  | 성경(504)  | 복지(858)  | 복지(2740)  | 사회(2923)  |
| 3  | 사람(58)   | 예수(400)  | 예수(494)  | 교육(759)  | 교육(2579)  | 복지(2460)  |
| 4  | 기독교(46)  | 기독교(389) | 신학(428)  | 특수(328)  | 특수(1295)  | 기출(1195)  |
| 5  | 하나님(46)  | 성경(364)  | 하나님(426) | 교육학(289) | 교육학(1211) | 교육학(1177) |
| 6  | 영성(39)   | 교회(330)  | 기도(307)  | 연구(264)  | 기출(1033)  | 장애(1120)  |
| 7  | 그리스(39)  | 그리스(293) | 그리스(287) | 장애(261)  | 장애(908)   | 특수(1099)  |
| 8  | 희망(39)   | 영성(278)  | 교회(282)  | 개론(257)  | 이해(891)   | 이해(988)   |
| 9  | 마음(37)   | 주석(258)  | 주석(272)  | 경제(251)  | 행정법(823)  | 헌법(966)   |
| 10 | 사랑(36)   | 기도(235)  | 영성(271)  | 이해(247)  | 개론(784)   | 장애인(932)  |
| 11 | 성경(34)   | 한국(220)  | 마음(256)  | 아동(247)  | 장애인(742)  | 행정법(921)  |
| 12 | 종교(30)   | 사랑(185)  | 종교(243)  | 이론(213)  | 분석(722)   | 상담(862)   |
| 13 | 신학(27)   | 신앙(172)  | 신화(222)  | 경매(213)  | 한국(707)   | 한국(836)   |
| 14 | 교회(27)   | 복음(169)  | 사랑(219)  | 이야기(189) | 경제(706)   | 투자(771)   |
| 15 | 당신(25)   | 마음(165)  | 역사(206)  | 실천(183)  | 대비(705)   | 연구(754)   |

변화정도를 종교와 사회과학 분야에서 이용된 자료의 표제에 기반한 관심 키워드의 출현빈도에 따른 순위를 보여 주고 있다.

조사 대상 기간 동안 이용된 대체자료의 표제로부터 추출된 종교분야의 이용자 관심 키워드의 변화는 크지 않은 것으로 나타났으며, 이

를 내용적으로 보면 기독교와 관련된 키워드가 대부분이었다. 사회과학 분야에서는 'A', 'B', 'C' 기간 공히 '사회복지', '특수 교육' 등과 관련된 키워드가 상위에 랭크되어 있다. 특히 '기출', '개론', '헌법' 등의 키워드가 'B', 'C' 기간에서 상위를 차지하고 있으며, 이는 사회적 현실이 반영되어 일반인들과 마찬가지로 장애인들이 관심을 가지는 자격증 또는 공무원 시험과 관련된 도서에 대한 관심이 증가했기 때문인 것으로 예측할 수 있다.

다음으로 구간으로 구분된 3개 기간의 변화에 따른 이용 대체자료 10개 주제 분야 가운데 자연과학(순수과학, 400)과 기술과학(500) 분야 대체자료의 키워드 변화 정도를 분석하였다. <표 15>는 기간별로 대체자료 이용 건수와 분야별 이용율의 변화정도를 자연과학과 기술과학 분야에서 이용된 자료의 표제에 기반한 관

심 키워드의 출현빈도에 따른 순위를 보여 주고 있다.

자연과학의 경우 이용 내역에서 데이터 건수가 적어 명확한 키워드 분석에 어려움이 있었을 뿐만 아니라 신뢰성에도 문제가 야기될 수도 있다. 다만 '우주', '행성'에 대한 관심이 지속적으로 존재하는 것으로 보이며, 시간의 흐름에 따라 '수학', '통계' 등의 키워드가 상위에 나타나는 것을 확인할 수 있다. 기술과학의 경우에도 이용 내역에서 차지하는 데이터 건수가 적어 명확한 키워드 분석이 어려우나 '침구', '침구학', '총통침법', '한의' 등 건강과 치료와 관련된 키워드가 지속적으로 등장하고 있다. 이러한 침술에 대한 관심이 꾸준히 존재하는 것으로 파악되었는데 이러한 현상은 장애인들의 직업선택과 관련이 깊을 것으로 예측할 수 있다.

<표 15> 기간별 자연과학 및 기술과학 분야의 대체자료 키워드 분석결과

|    | 자연과학        |          |          | 기술과학     |          |           |
|----|-------------|----------|----------|----------|----------|-----------|
|    | A           | B        | C        | A        | B        | C         |
| 1  | 과학(158)     | 생태(521)  | 과학(426)  | 치료(285)  | 치료(1517) | 치료(2014)  |
| 2  | 지구(40)      | 이슈(516)  | 이야기(156) | 건강(205)  | 총통(938)  | 총통(1399)  |
| 3  | 최신(35)      | 돌보기(516) | 수학(146)  | 임상(186)  | 건강(803)  | 통증(1167)  |
| 4  | 해양(35)      | 우화(516)  | 우주(114)  | 침구(124)  | 통증(684)  | 건강(845)   |
| 5  | 행성(32)      | 과학(459)  | 통계(96)   | 침구학(110) | 의학(680)  | 의학(796)   |
| 6  | 입문(30)      | 이야기(192) | 인간(85)   | 의학(100)  | 임상(605)  | 근육(775)   |
| 7  | ENGLISH(29) | 수학(162)  | 지구(80)   | 평형(92)   | 침구(597)  | 임상(672)   |
| 8  | 이야기(25)     | 우주(141)  | 동물(71)   | 통증(91)   | 동의(545)  | 인체(595)   |
| 9  | 통계(25)      | 지구(96)   | 시간(64)   | 요법(78)   | 요법(453)  | 체질(595)   |
| 10 | 수학(23)      | 세상(82)   | 세상(59)   | 아이(75)   | 해부학(401) | 요법(538)   |
| 11 | GRAMMAR(22) | 우리(82)   | 생명(58)   | 우리(71)   | 근육(391)  | 침구(497)   |
| 12 | 환경(22)      | 물리학(78)  | 공부(58)   | 기술(71)   | 한의(385)  | 총통침법(445) |
| 13 | 파놉티콘(22)    | 화학(69)   | 진화(56)   | 생명(67)   | 보감(357)  | 강의록(445)  |
| 14 | 지질(21)      | 인간(68)   | 양자(55)   | 백과(67)   | 강의록(349) | 동의(416)   |
| 15 | NORTH(19)   | 동물(68)   | 코스모스(53) | 문원(66)   | 사암(346)  | 아이(391)   |

다음으로 구간으로 구분된 3개 기간의 변화에 따른 이용 대체자료 10개 주제 분야 가운데 예술(600)과 언어(700) 분야 대체자료의 키워드 변화 정도를 분석하였다. <표 16>은 기간별로 대체자료 이용 건수와 분야별 이용율의 변화 정도를 예술과 언어 분야에서 이용된 자료의 표제에 기반한 관심 키워드의 출현빈도에 따른 순위를 보여준다.

예술의 경우 '피아노', '음악'과 관련된 키워드가 꾸준히 상위를 차지하고 있다. 'A' 기간에서는 '피아노'에 대한 키워드가 주를 이루고 있으며, 주로 음악과 관련된 키워드가 상위 20위를 차지하였다. 그러나 'B'와 'C' 기간으로 이동하면서 점차 '드라마', '그림', '축구', '스포츠' 등과 같은 키워드가 상위에 나타나고 있음을 보여준다. 이는 이용자들의 음악 분야뿐만 아니라 예술분야 전반에 대한 관심이 증가한 것으

로 해석할 수도 있다. 다만 국립장애인도서관의 예술 분야 대체자료가 다루는 주제 범위의 변화와 대체자료 제작 추세가 어떻게 변화했는지를 확인할 필요가 있다.

언어 분야에서는 '영어', '국어'가 꾸준히 최상위로 랭크되어 있다. 또한 '해커스', '재정', '공무원', '선재', '기출', '교육부'와 같은 키워드가 상위에 랭크되어 있는 것으로 보아 이 시기 공무원이나 자격시험과 같은 수험서에 대한 관심이 크게 증가한 것으로 해석할 수 있을 것이다. 'A' 기간에서는 '단어', '암기', 'GRAMMAR' 등이 상위에 존재하는 반면 'B'와 'C' 기간에는 공무원 수험서에서 추출된 키워드가 주를 이루고 있으며, 특히 '선재국어', '재정국어', '교육부 지정' 등에서 추출된 것으로 보이는 키워드가 다수 존재하는 것으로 나타났다.

다음의 <표 17>은 3개로 구분된 기간 동안의

<표 16> 기간별 예술 및 언어 분야의 대체자료 키워드 분석결과

|    | 예술           |               |           | 언어           |          |          |
|----|--------------|---------------|-----------|--------------|----------|----------|
|    | A            | B             | C         | A            | B        | C        |
| 1  | 피아노(526)     | 피아노(816)      | 음악(258)   | 영어(328)      | 영어(2239) | 영어(3641) |
| 2  | 성악(239)      | PIANO(607)    | 영화(168)   | 국어(143)      | 국어(1080) | 국어(1211) |
| 3  | PIANO(119)   | 작품집(540)      | 클래식(159)  | ENGLISH(124) | 기출(618)  | 단어(933)  |
| 4  | MAJOR(118)   | 음악(341)       | 게임(103)   | 스피드(123)     | 해커스(585) | 암기(646)  |
| 5  | 음악(118)      | MAJOR(227)    | PIANO(99) | 암기(123)      | 단어(494)  | 스피드(645) |
| 6  | 바이올린(111)    | SONATAS(227)  | 한국(93)    | 자동(105)      | 공무원(460) | 원어(630)  |
| 7  | MAZURKEN(98) | 드라마(213)      | 사전(92)    | 원어(105)      | 일본어(379) | 해커스(625) |
| 8  | CONCERTO(82) | 시대(197)       | 문화(78)    | 단어(99)       | 암기(369)  | 자동(614)  |
| 9  | VIOLIN(81)   | SOLO(191)     | 피아노(75)   | 단계(98)       | 스피드(361) | 단계(595)  |
| 10 | 이론서(76)      | 한국(188)       | 그림(75)    | 시험(86)       | 기초(359)  | 공무원(552) |
| 11 | 교본(75)       | SBS(181)      | 볼링(73)    | 토익(84)       | 시험(327)  | 필수(506)  |
| 12 | CHANSONS(75) | 야인(181)       | 이야기(70)   | 일본어(80)      | 단계(315)  | 회화(452)  |
| 13 | SONATA(74)   | 대하(181)       | 예술(70)    | 해커스(73)      | 선재(314)  | 영문법(433) |
| 14 | LODIES(71)   | SONATA(180)   | 시대(68)    | 공무원(71)      | 원어(309)  | 선재(407)  |
| 15 | 어드벤처(71)     | CONCERTO(176) | 사랑(68)    | 한국어(69)      | 자동(304)  | 기출(403)  |

〈표 17〉 기간별 문학 및 역사 분야의 대체자료 키워드 분석결과

|    | 문학       |           |           | 역사       |           |           |
|----|----------|-----------|-----------|----------|-----------|-----------|
|    | A        | B         | C         | A        | B         | C         |
| 1  | 소설(335)  | 소설(7992)  | 소설(13251) | 한국사(350) | 한국사(1262) | 한국사(1663) |
| 2  | 이야기(245) | 장편(5452)  | 장편(9132)  | 이야기(184) | 이야기(872)  | 역사(789)   |
| 3  | 한국(171)  | 완결(3795)  | 완결(7862)  | 조선(169)  | 역사(790)   | 조선(521)   |
| 4  | 아버지(151) | 판타지(3668) | 판타지(7093) | 역사(141)  | 조선(545)   | 세계사(498)  |
| 5  | 장편(149)  | 무협(3169)  | 무협(4809)  | 중국(136)  | 한국(420)   | 이야기(475)  |
| 6  | 단편(130)  | 사랑(856)   | 현대(1512)  | 고종(126)  | 세계(403)   | 한국(407)   |
| 7  | 소나기(113) | 현대(786)   | 사랑(1398)  | 시험(99)   | 세계사(365)  | 기출(328)   |
| 8  | 세상(112)  | 문학(749)   | 무적(1106)  | 검정(99)   | 여행(271)   | 여행(309)   |
| 9  | 운수(108)  | 이야기(677)  | 전기(1036)  | 능력(99)   | 기출(242)   | 시대(287)   |
| 10 | 봄봄(105)  | 아이(540)   | 마법사(983)  | 한국인(82)  | 세상(194)   | 세계(278)   |
| 11 | 7년후(99)  | 사람(443)   | 이야기(901)  | 세계사(82)  | 전쟁(185)   | 공무원(277)  |
| 12 | 세계(95)   | 전설(435)   | 무림(858)   | 대비(72)   | 국사(183)   | 영식(251)   |
| 13 | 또래(95)   | 한국(435)   | 사람(787)   | 기분(72)   | 중국(175)   | 해동(224)   |
| 14 | 문학(91)   | 우리(415)   | 전쟁(762)   | 고급(72)   | 문화(166)   | 실록(210)   |
| 15 | 시집(86)   | 전기(397)   | 황금(746)   | 세상(64)   | 왕조(163)   | 전쟁(208)   |

변화에 따른 이용 대체자료 10개 주제 분야 가운데 문학(800)과 역사(900) 분야 대체자료의 키워드 변화 정도를 분석하였다. 〈표 17〉에서 기간별로 대체자료 이용 건수와 분야별 이용율의 변화정도에 나타난 바와 같이 예술과 언어 분야에서 이용된 자료의 표제에 기반한 관심 키워드의 출현빈도에 따른 순위이다.

문학의 경우 표제에 나타난 키워드를 통해 관심사를 파악하기에 한계가 있는 주제 분야이긴 하지만 지속적으로 판타지 소설에 대한 관심이 존재하는 것으로 분석 결과 나타났다. ‘소설’이라는 용어가 조사된 3개 기간 동안 지속적으로 1위에 랭크되어 있으며, ‘장편’, ‘판타지’, ‘무협’ 등이 상위 집단으로 나타나고 있다. 특히 ‘A’ 기간 동안에는 ‘아버지’가 4위에 올라와 있고 ‘B’와 ‘C’ 기간에는 ‘사랑’, ‘사람’, ‘행복’ 등이 상위에 집중되는 현상을 보이고 있다.

역사 분야에서는 ‘한국사’, ‘중국’, ‘역사’와 같은 키워드가 상위에 집중적으로 등장하고 있으며, 또한 ‘A’ 기간에서는 ‘한국사’와 함께 ‘능력’, ‘검정’, ‘대비’, ‘고급’ 등과 같은 키워드가 상위에 나타나는데 이를 통해 ‘한국사 능력검정시험’에 대한 수험서가 상당히 많이 이용되었음을 알 수 있다. ‘B’ 기간에서는 ‘세계사’와 ‘조선 왕조실록’에 대한 키워드가 상위로 나타나는 것을 확인할 수 있으며, C기간에서는 ‘공무원’, ‘해동’, ‘설민석’과 같은 키워드가 나타나고 있어 사회과학이나 언어에서와 마찬가지로 공무원 시험에 대한 관심이 증가한 것으로 예측할 수 있다.

지금까지 DREAM을 통하여 다운로드된 대체자료의 표제로부터 추출한 키워드를 활용하여 장애인들이 주제 분야별로 조사 대상 기간 동안에 어떠한 자료를 많이 이용하였는가를 분

석하였다. 대체자료 표제에 기반한 키워드 분석을 통하여 시대적, 환경적 변화에 따라 사회적 관심사가 변화되고 이용자들의 관심사와 정보 요구에 대한 변화를 개략적으로 예측해 볼 수는 있지만 표제 외의 문헌의 다양한 속성을 활용하지 않았기 때문에 분석 결과에 대한 객관성에는 한계를 가질 수 있다. 따라서 단순히 표제로부터의 키워드뿐만 아니라 초록이나 요약문, 나아가 본문의 내용을 활용할 수 있다면 분석결과에 대한 신뢰성을 확보할 수 있을 것이다.

#### 4.3 국립장애인도서관 소장 자료 이용 내역 분석

이용자 선호도 분석을 위해 사용한 데이터는 DREAM의 이용 내역 데이터로, 해당 데이터에는 DREAM에 참여하고 있는 약 30여개 기관의 대체자료 이용 내역을 포함하고 있다. 이 가운데 장애인 이용자의 정보 요구에 부응하는 대체자료 개발을 위해 국립장애인도서관 소장 대체자료로 국한하여 이용자의 대출 이력 데이터를 바탕으로 이용자 선호도를 분석하였다. 이를 위하여 DREAM의 이용 내역 데이터에서 소장도서관이 '국립중앙도서관'인 대체자료의 이용 내역을 추출하여 '국립중앙도서관 대체자료 이용내역' 데이터로 재구성하였다. 재구성된 데이터 중에서 이용자 정보가 기재되어 있는 이용 내역을 다시 추출하여 총 96,475건의 이용 내역을 이용자 선호도 분석에 사용하였다.

이용자 선호 주제 분석에 앞서 보다 정확한 분석 결과를 도출하기 위해 DREAM을 통하여 서비스되고 있는 국립중앙도서관 대체자료의

주제 구성을 분석하였다. 현재 DREAM을 통해 서비스 되고 있는 국립장애인도서관 대체자료는 총 121,326건이며, 그 중 사회과학 주제 분야의 대체자료가 32,730건(27.0%)으로 가장 많이 제작되어 서비스되고 있으며, 기술과학, 문학, 총류 분야에서 각각 17,698건(14.6%), 17,687건(14.6%), 17,201건(14.2%)의 대체자료가 서비스 되고 있다. 또한 가장 적은 대체자료가 제작되어 서비스되고 있는 주제 분야는 자연과학으로 전체의 1.4%에 해당하는 1,678건의 대체자료가 서비스되고 있다. 대체자료의 주제별 구성비율은 문학, 사회과학, 기술과학으로 이어지는 실제 DREAM에서 조사된 이용율은 차이를 보이고 있다.

반면, 국립중앙도서관에 소장된 대체자료 96,475건을 대상으로 이용자 선호 주제 분야를 분석한 결과 문학 분야의 대체자료 이용 건수가 전체 이용 건수 380,565건의 42.1%에 해당하는 160,032건으로 가장 높게 나타났으며, 다음으로 사회과학 주제 분야의 대체자료 이용 건수가 72,746건(19.1%), 기술과학과 철학 분야의 대체자료 이용 건수가 각각 36,102건(9.5%), 28,667건(7.5%)으로 다음으로 높게 나타났다. 가장 낮은 선호도를 보이는 주제 분야는 자연과학으로 전체의 1.5%의 해당하는 5,893건을 차지하였다. 전반적으로 국립중앙도서관 대체자료의 주제 구성 비율과 DREAM 이용자 선호도에는 차이를 보이고 있는데, 특히 문학의 경우 대체자료 구성 비율은 10개 주제 분야 가운데 3위(14.6%)에 불과하지만 실제 이용율에 있어서는 42.1%를 차지하였다. 반면 총류의 경우 대체자료 구성 비율이 2위인 14.6%에 이르지 않지만 실제 이용율 순위에서는 8위인 3.5%에 머

물러 실제 이용율과 대체자료 구성 비율에는 큰 차이를 보이고 있다.

다음으로 DREAM의 대체자료 전체 이용 내역을 대상으로 이용자의 관심키워드를 도출한 결과는 <표 18>과 같다. 먼저 <표 18>은 문학을 제외한 9개 주제 분야에 대한 키워드 도출 결과를 TF-IDF 가중치에 따라 정렬하여, 상위 10개의 키워드와 해당 키워드의 TF-IDF 수치를 추출한 내용이다.

전체 이용자의 이용내역 데이터를 통해 이용자 관심 키워드 분석 결과 총류에서는 '프로그래밍', '안드로이드', '자바', '평설', '독서'와 같은 키워드가 상위에 도출되었는데 이는 정보기술과 컴퓨터와 관련된 도서에 대한 이용률이 높기 때문인 것으로 보인다. 또한 '파이', '스크

립트', '딥러닝'과 같은 키워드가 도출되었으며, 이를 통해 컴퓨터 공학과 관련된 이용자 관심이 높은 것을 확인할 수 있다. 철학의 경우 '명리', '상담', '사주'와 같은 키워드가 도출되었으며, 종교 분야에서는 '성경', '신학', '교회', '기독교', '예수'와 같은 키워드가 도출되어 기독교 관련 서적에 대한 관심이 높은 것을 확인할 수 있다.

사회과학 분야의 경우 '행정법', '부동산', '헌법', '행정법', '특수교육' 등과 같은 키워드가 도출되었으며, 이를 통해 부동산이나, 특수 교육과 관련된 도서에 대한 관심이 높은 것으로 나타났다. 자연과학 분야의 경우 '과학', '식물', '물리학', '태양계', '생명'과 같은 키워드가 도출되었으며, 기술과학 분야에서는 '통증', '근육',

<표 18> DREAM 이용자의 주제 분야별 관심 키워드(문학 제외)

| 순위 | 총류       | 철학       | 종교      | 사회과학     | 자연과학    | 기술과학     | 예술       | 언어       | 역사       |
|----|----------|----------|---------|----------|---------|----------|----------|----------|----------|
| 1  | 프로그래밍    | 명리       | 성경      | 행정법      | 과학동아    | 통증       | 뽀롱뽀롱     | 영단어      | 한국사      |
|    | 1082.703 | 1872.576 | 970.075 | 3227.723 | 140.622 | 2511.809 | 597.645  | 2012.657 | 2256.819 |
| 2  | 안드로이드    | 사주       | 신학      | 부동산      | 식물      | 체질       | 야인시대     | 원어민      | 해동       |
|    | 1032.695 | 1354.589 | 808.579 | 2175.252 | 127.847 | 1645.721 | 457.0227 | 1973.107 | 624.012  |
| 3  | 자바       | 사주명리학    | 기독교     | 헌법       | 코스모스    | 근육       | 객석       | 원어       | 영식       |
|    | 786.606  | 709.704  | 641.446 | 2033.513 | 123.334 | 1630.420 | 432.853  | 1935.754 | 551.503  |
| 4  | 평설       | 상담       | 하나님     | 행정법총론    | 물리학     | 총통       | 뽀로       | 영어자동암기   | 문화유산답사   |
|    | 783.624  | 488.585  | 604.954 | 1968.713 | 120.847 | 1450.168 | 409.109  | 1891.810 | 531.728  |
| 5  | 독서       | 육효       | 교회      | 행정학      | 통계      | 총통침법     | 클래식      | 해커스      | 국사       |
|    | 772.816  | 443.839  | 594.349 | 1579.281 | 96.501  | 1313.940 | 258.687  | 1729.216 | 529.531  |
| 6  | 파이       | 미움       | 주석      | 교육학      | 태양계     | 치료       | 피아노      | 국어       | 고종훈      |
|    | 542.527  | 384.514  | 516.348 | 1550.499 | 81.297  | 1265.457 | 219.722  | 1569.150 | 516.348  |
| 7  | 드로       | 타로       | 실교      | 특수교육학    | 수학      | 침구       | 음반       | 초스피드     | 실민석      |
|    | 520.742  | 382.036  | 410.881 | 1278.785 | 73.132  | 1081.034 | 210.934  | 1404.808 | 443.839  |
| 8  | 자바스크립트   | 박청화      | 기도      | 경매       | 과학사     | 지압       | 음악       | 영문법      | 해동한국사    |
|    | 476.797  | 680.120  | 397.356 | 1201.882 | 72.508  | 1034.893 | 200.300  | 1344.701 | 431.078  |
| 9  | 스크립트     | 명리학      | 영성      | 복지       | 상대성     | 도수       | 르트       | 선재국어     | 왕조       |
|    | 448.215  | 369.134  | 379.515 | 1130.664 | 68.114  | 1021.709 | 173.580  | 1316.137 | 366.937  |
| 10 | 딥러닝      | 사주명리     | 신앙      | 신용       | 생명과학    | 마사지      | 유성기      | 자동차      | 통감       |
|    | 404.289  | 340.359  | 360.979 | 1017.315 | 64.675  | 1006.329 | 162.595  | 1305.539 | 360.345  |

‘치료’, ‘침법’, ‘총통’, ‘체질’과 같은 키워드가 도출되어 건강이나 통증치료에 대한 이용자의 관심도가 높은 것으로 나타났다.

예술 분야에서는 ‘객석’, ‘음반’과 같은 음악 관련 키워드와 ‘야인시대’, ‘뽀로’와 같은 키워드가 도출되었다. 언어분야에서는 ‘영단어’, ‘원어’, ‘해커스’, ‘국어’, ‘문제집’과 같이 공무원 시험서와 관련된 키워드들이 도출되었다. 마지막으로 역사 분야에서는 ‘한국사’, ‘역사’, ‘국사’ 등 우리 역사와 관련된 키워드가 주로 도출되었다.

국립대체자료공유시스템인 DREAM을 통한 대체자료 이용내역에 기반한 주제 분야별 이용자 선호도와 국립장애인도서관 소장자료에 대한 이용자 선호도 및 이용내역을 비교 분석한 결과, 총류를 제외한 나머지 주제 분야의 대체자료에 대한 이용자 선호도는 차이가 없는 것으로 나타났다. 총류의 경우 국립장애인도서관 소장자료에 대한 이용자들의 선호도에서 ‘A’ 기간에는 ‘피아노’, ‘작품집’이 ‘B’와 ‘C’ 기간에는 ‘프로그래밍’과 ‘독서’가 최상위에 랭크된 것과 비교해 볼 때 DREAM을 통한 대체자료 이용내역을 통한 분석결과에는 차이를 보이고 있다. 특히 대체자료 이용시기별로 ‘A’(2013.03.07. ~ 2015.12.31.), ‘B’(2016.01.01. ~ 2017.10.31.), 그리고 ‘C’(2017.11.07. ~ 2019.09.30.)로 구분하였을 때, 시기에 따라 이용자들의 관심 분야에는 차이가 있는 것으로 나타났다.

문학 분야의 도서의 경우 표제가 본문의 내용을 대표하지 않은 경우가 다른 주제 분야에 비하여 빈번하게 발생하고 있기 때문에 표제에 기반한 키워드의 추출과 이를 통한 이용자의 관심도를 측정하기에는 제약이 따른다. Open

API 검색을 통해 문학 분야 도서의 카테고리 정보를 수집하여 관심 키워드를 도출하였으며, 그 결과는 다음의 <표 19>와 같다.

<표 19> 전체 이용자의 주제별 관심 키워드 (문학)

| 순위 | 키워드(카테고리 명)           |
|----|-----------------------|
| 1  | 잡지   문예/교양지   문예      |
| 2  | 시/에세이                 |
| 3  | 어린이   어린이(공통)   어린이동화 |
| 4  | 소설   나라별소설   한국소설     |
| 5  | 소설                    |
| 6  | 어린이                   |
| 7  | 시/에세이   한국시           |
| 8  | 소설   장르소설   SF/판타지    |
| 9  | 소설   나라별소설   영미소설     |
| 10 | 인문   문학론   문학이론       |

전체 이용자의 문학 주제 분야 관심도는 문예 잡지에서 가장 높게 나타났으며, 이는 장애인 문학지인 ‘숫대문학’의 이용률이 높기 때문인 것으로 추측된다. 다음으로 관심도가 높은 키워드는 ‘시/에세이’이며, ‘어린이’, ‘동화책’, ‘한국소설’에 대한 관심도가 높은 것으로 나타났다.

이 연구에서는 국립중앙도서관의 Open API를 활용하여 장애인 이용자를 위한 대체자료 제작 대상 리스트를 작성할 수 있도록 지원하는 방법을 제안하였다. Open API 검색을 위하여 이용자 대상의 주제별 관심 키워드를 질의어로 활용할 수 있으며, 또한 검색결과를 일치도 순으로 정렬하여 최상위 도서, 즉 대체자료 제작 대상 우선순위에 대한 정보의 수집이 가능하다. 물론 대체자료 제작 대상 목록은 주제별로 도서 정보를 포함할 수 있으며, 검색결과



를 키워드와의 일치도 순으로 정렬이 가능하기 때문에 해당 키워드와 동일한 표제를 가지는 도서 정보의 수집이 가능하다.

#### 4.4 외부 서비스 및 정보원과의 연계방안

이용자 선호도를 반영한 대체자료 제작대상 도서를 선정하기 위해서는 다양한 외부 서비스 및 정보 원천과의 연계를 통해 보다 다양하고 방대한 이용자 정보 및 도서 정보 데이터를 수집해야 한다. 이를 위하여 <그림 3>은 다양한 외부 서비스 및 정보원과의 연계 방안을 도식화한 것이다.

먼저 공공도서관과의 연계를 통해 이용자 정보 및 도서 정보를 수집할 수 있으며, 대형 서점 및 포털과의 연계를 통해 베스트셀러 정보나 권장/추천 도서 정보를 수집할 수 있다. 또한 다양한 미디어 기관과의 연계를 통해 장애인 관련 이슈, 정책동향, 신문기사와 같은 동향 정보를 수집한다. 이와 같이 다양한 정보원과의 연계를 통해 데이터가 수집되면, 국립장애인도서관은

장애인 이용자의 대출이력 데이터를 활용하여 이용자의 요구를 반영한 대체자료 제작 대상 도서를 선정한다. 국립장애인도서관은 해당 도서들을 직접 대체자료로 제작하여 DREAM과 같은 대체자료 공유 서비스를 통해 다양한 공공도서관의 이용자에게 제공이 가능하다. 또한 대체자료 제작 추천 리스트를 연계 기관에 제공하여, 연계 기관의 대체자료 장서 확충에 기여할 수도 있다.

## 5. 결 론

이 연구의 목적은 국립장애인도서관의 이용자 데이터 및 이용 내역 데이터를 기반으로 이용자의 선호 주제 분야 및 관심 키워드를 분석하고, 이용자 관심 정보를 반영할 수 있는 도서자동분류엔진의 설계와 장애인들이 대체자료를 선택함에 있어서 개인에게 적합하고, 필요한 도서를 합리적으로 제공할 수 있도록 이용자 선호도 기반 도서추천시스템을 시범 구축하는 것이다.



<그림 3> 외부 서비스 및 정보원과의 연계 방안

국립장애인도서관의 이용자 선호도 및 관심 분야 정보를 수집하기 위하여 웹상에서 이루어지는 이용자의 다양한 활동 기록을 수집하고, 이를 통해 이용자 프로파일을 구축하는 기법인 웹 기반의 동적 프로파일링 모델을 활용하였으며, DREAM을 통한 대체자료의 다운로드 내역을 활용하여 이용자의 선호 주제 분야 및 주제 분야별 관심 키워드, 그리고 이용자 선호도를 분석하였다.

대출정보를 이용한 이용자 선호 주제 분야 및 주제 분야별 이용자 선호도 분석은 통계분석방법을 통해 이루어졌으며, 이용자 관심 키워드의 추출은 대체자료 이용내역 데이터를 기반으로 형태소 분석 및 Open API를 통해 수행되었다. 추출된 키워드에 대한 TF-IDF 가중치를 계산하여 키워드의 중요도를 도출하였으며, 또한 이용자 선호도를 반영하여 대체자료 제작 대상 도서를 선정하기 위한 방법론을 제안하였다.

수집 대상은 2013.03.07일부터 2019.09.30일까지 약 6년 6개월 동안의 DREAM시스템의 대체자료 다운로드 이력 데이터로 총 568,945이다. 이 가운데 대체자료명이나 분류번호, 그리고 대체자료명에 명사를 포함하고 있지 않은 98,511건이 전처리 과정을 거쳐 제거되어 이용내역과 대체자료 다운로드 이력 분석을 위하여 사용된 데이터는 총 470,434건이다.

이용자 관심 키워드 분석은 DREAM 이용내역 데이터를 10개 주제에 따라 구분하고, 각 주제에 해당하는 이용내역 개별 데이터에서 키워드를 추출하여 각 키워드별 중요도를 계산하는 방법으로 수행되었다. 기간에 따른 주제별 대체자료 이용율은 'A' 기간의 경우 총류, 사회과학, 문학의 순으로 높게 나타났으나 언

어, 종교, 그리고 자연과학의 이용율은 낮게 나타났다. 'B' 기간에서는 문학, 사회과학, 총류, 기술과학의 순으로 나타나 'A' 기간과 비교하여 문학에 대한 이용율이 가장 높게, 그리고 사회과학과 총류가 그 뒤를 이어 이용자들의 관심과 이용율의 변화가 일어나고 있음을 보여준다. 마지막으로 'C' 기간에서의 이용율은 문학, 사회과학, 기술과학, 철학의 순으로 나타나 'A' 기간과 비교하여 총류에 대한 이용율은 감소한 반면에 문학, 사회과학, 그리고 기술과학은 증가한 것으로 나타났다.

다음으로 DREAM의 이용내역 데이터 가운데 국립장애인도서관 소장 대체자료에 대한 이용자의 대출 이력 데이터 96,745건을 이용하여 이용자 선호도를 분석한 결과, 사회과학 주제 분야의 대체자료 이용 건수가 26,377건(27.3%)로 가장 높게 나타났으며, 기술과학과 총류 분야의 대체자료 이용 건수가 각각 15,296건(15.9%), 15,181건(15.7%)으로 다음으로 높게 나타났다. 문학의 경우 소장 건수에 비해 이용율이 기술과학이나 총류에 비해 상대적으로 낮은 것으로 조사되었다. 가장 낮은 선호도를 보이는 주제 분야는 자연과학으로 전체의 1.3%의 해당하는 1,210건을 차지하였으며, 전반적으로 국립중앙도서관 대체자료의 주제 구성 비율과 차이가 없는 것으로 나타났다.

이 연구에서 대체자료 제작 대상 도서의 선정을 위하여 실험적으로 수행한 것이기 때문에 활용도에 한계가 있을 수 있다. 이러한 한계를 극복하기 위해서는 이용자들의 도서 이용률 및 관심도가 반영된 도서관의 다대출 도서나 도서관별 추천도서 그리고 대형 서점의 베스트셀러를 대상으로 이용자 선호도에 대한 주요 키워

드를 반영하여 대체자료 제작 대상 도서를 선정하는 것이 중요하다. 정확도 및 활용도가 높은 대체자료 제작 대상 도서 리스트를 선별하기 위해 다양한 연계기관들과의 서비스 및 정

보 연계가 요구되며, 이를 지원하기 위해 다양한 외부 서비스 및 정보원과의 연계방안이 필요하다.

## 참 고 문 헌

- [1] 강성구, 임경원. 2017. 장애인용 대체자료 제작 전략 및 환경 변화에 관한 연구. 『한국도서관·정보학회지』, 48(4): 283-301.
- [2] 국립장애인도서관. 2017. 『빅데이터 분석을 통한 대체자료 서비스 전략 수립 연구』. 서울: 국립장애인도서관.
- [3] 국립장애인도서관. 2019. 『2019 도서관 장애인서비스 현황조사』. 서울: 국립장애인도서관.
- [4] 권혁인 외. 2015. 도서정보 기반의 고객 맞춤형 큐레이션 서비스 및 비즈니스 모델 연구. 『한국IT서비스학회지』, 14(1): 251-262.
- [5] 김미지 외. 2018. 키워드 가중치를 이용한 뉴스 기사에서의 이슈 키워드 자동 추출 시스템. 『한국방송·미디어공학회 추계학술대회』, 146-148.
- [6] 김수아, 조희선, 이현아. 2015. 다양한 어휘 가중치를 이용한 블로그 포스트의 자동 분류. 『한국마린엔지니어링학회지』, 39(1): 58-62.
- [7] 대통령소속 도서관정보정책위원회. 2019. 『제3차 도서관발전종합계획: 2019-2023』. 서울: 대통령소속 도서관정보정책위원회.
- [8] 백재은. 2016. 장애인의 독서치료를 위한 대체자료 현황 및 개발방안에 관한 연구. 『한국문헌정보학회지』, 50(4): 165-184.
- [9] 안희정, 김기원, 김승훈. 2017. 도서 정보 및 본문 텍스트 통합 마이닝 기반 사용자 맞춤형 도서 큐레이션 시스템. 『정보기술응용연구』, 24(1): 33-43.
- [10] 오선경. 국내 시각장애인도서관의 대체자료 제작현황 및 개선방안 연구. 『한국문헌정보학회지』, 51(3): 215-246.
- [11] 윤성희. 2016. 질의기반 사용자 프로파일을 이용하는 개인화 웹 검색. 『한국산학기술학회논문지』, 17(2): 690-696.
- [12] 윤희운. 2010. 주요 국가의 장애인용 대체자료 개발정책 분석. 『한국도서관·정보학회지』, 41(1): 29-49.
- [13] 이정미. 2013. 빅데이터의 이해와 도서관 정보서비스에의 활용. 『한국비블리아학회지』, 24(4): 53-73.

- [14] 장보성, 서만덕, 남영준. 2012. 우리나라 도서관의 장애인서비스에 관한 분석 연구. 『한국비블리아학회지』, 23(2): 87-107.
- [15] 조현양. 2017. 자동분류기반 성격 유형별 도서추천시스템 개발을 위한 실험적 연구. 『한국도서관·정보학회지』, 48(2): 215-236.
- [16] 한희준, 최윤수, 최성필. 2018. 개인 관심분야 추적기법을 이용한 과학기술정보 개인화에 관한 연구. 『한국문헌정보학회지』, 52(3): 5-33.
- [17] Mikolov, T. et al. 2013. "Efficient estimation of word representations in vector space." arXiv preprint arXiv: 1301.3781. pp. 1-12. <<https://arxiv.org/pdf/1301.3781.pdf>>
- [18] Mikolov, T., Yih, W. T. and Zweig, G. 2013, June. "Linguistic regularities in continuous space word representations." In *hlt-Naacl*, vol. 13, pp. 746-751.
- [19] 국가대체자료공유시스템 DREAM. [online] [cited 2020. 7. 14.] <<http://dream.nl.go.kr/>>

• 국문 참고자료의 영어 표기

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

- [1] Kang, seong-goo and Lim, Kyoung-won. 2018. "A Study on the Change of Production Strategy and Environment for Alternative Material for Person with Special Needs." *Journal of the Korean Library and Information Science Society*, 48(4): 283-301.
- [2] National Library for the Disabled. 2017. *The Study on the Development of Strategic Service Planning for Alternative Materials Using Big Data Analysis*. Seoul: National Library for the Disabled.
- [3] National Library for the Disabled. 2019. *2019 The Investigation on Current State of Library Services for the Disabled*. Seoul: National Library for the Disabled.
- [4] Kwon, Hyeog-In et al. 2015. "A Study of a Personalized Curation Services and Business Model based on Book Information." *Journal of Information Technology Services*, 14(1): 251-262.
- [5] Kim, M. J. et al. 2018. "Automatic Keyword Extraction in News Articles for Trend Tracking." *Proceedings of the Korean Society of Broadcast Engineers Conference*, 146-148.
- [6] Kim, Suah, Jho, Hee-Sun and Lee, Hyun Ah. 2015. "Automatic Classification of Blog Posts Considering Category-specific Information." *Journal of the Korean Society of Marine Engineering*, 39(1): 58-62.
- [7] Presidential Committee on Library and Information Policy of Korea. 2019. *Comprehensive Plans for Libraries in Korea: 2019-2023*. Seoul: Presidential Committee on Library and Information Policy of Korea.

- [8] Baek, Jae-Eun. 2016. "The Study about Analyzing the Present Situation of Alternative Formats for Bibliotherapy of the Disabled People." *Journal of the Korean Society for Library and Information Science*, 50(4): 165-184.
- [9] Ahn, Hee-Jung, Kim, Kee-Won and Kim, Seung-Hoon. 2017. "Personalized Book Curation System based on Integrated Mining of Book Details and Body Text." *Journal of Information Technology Applications & Management*, 24(1): 33-43.
- [10] Oh, Seonkyung. 2017. "A Study on the Improvement and the Production State of Alternative Materials of Special Libraries for the Visually Disabled in Korea." *Journal of the Korean Society for Library and Information Science*, 51(3): 215-246.
- [11] Yoon, Sung Hee. 2016. "Personalized Web Search using Query based User Profile." *Journal of the Korea Academia-Industrial Cooperation Society*, 17(2): 690-696.
- [12] Yoon, Hee-Yoon. 2010. "Analysis of Alternative Formats Development Policy for the Disabled Persons in the Major Countries." *Journal of the Korean Library and Information Science Society*, 41(1): 29-49.
- [13] 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.
- [14] Jang, Bo-Seong, Suh, Mandeok, and Nam, Young-Joon. 2012. "The Analysis Research about Library Service for the Disabled in Korea." *Journal of the Korean Biblia Society for Library and Information Science*, 23(2): 87-107.
- [15] Cho, Hyun-Yang. 2017. "A Experimental Study on the Development of a Book Recommendation System Using Automatic Classification, Based on the Personality Type." *Journal of the Korean Library and Information Science Society*, 48(2): 215-236.
- [16] Han, Heejun, Choi, Yunsoo and Choi, Sung-Pil. 2018. "A Study on Personalization of Science and Technology Information by User Interest Tracking Technique." *Journal of the Korean Society for Library and Information Science*, 52(3): 5-33.
- [19] National Library for Disabled. DREAM - Direct Rapid Easy Accessible Material Service. [online] [cited 2020. 7. 14.] <<http://dream.nl.go.kr/>>

