Journal of Korean Library and Information Science Society
Vol. 54, No. 3

OPEN ACCESS

공공도서관 서지데이터의 품질 제고 방안*

- 부산시 공공도서관을 중심으로 -

Improving the Quality of Bibliographic Data in Public Libraries: Focusing on Public Libraries in Busan Metropolitan City

노 지 현(Jee-Hyun Rho)**
이 은 주(Eun-Ju Lee)***

〈 목 차 〉

Ⅰ. 서 론
Ⅱ. 서지데이터의 구축방식
Ⅲ. 서지데이터의 품질
Ⅳ. 서지데이터의 문제점과 품질 제고 방안
Ⅴ. 결 론

요 약 : 부산지역에서는 2020년 대표도서관을 중심으로 모든 도서관이 공동으로 사용할 통합도서관시스템(ILS)을 구축하고, 지역 내 49개 공공도서관과 103개 공립작은도서관의 서지데이터를 통합하였다. 그러나 각 도서관은 여전히 개별적으로 서지데이터를 구축하고 있고, 각 도서관에서 구축한 서지데이터는 물리적으로만 통합되어 있을 뿐 통합환경을 활용한 업무효율이나 데이터의 고품질화를 이루지 못하고 있다. 이러한 배경에서, 이 연구는 부산지역 공공도서관의 서지데이터 구축체계와 품질을 분석하고 통합환경에 적합한 새로운 실행전략을 제안하는데 목적을 두고 있다. 이를 위해 (1) 부산지역 공공도서관에서 서지데이터를 구축하는 방식을 조사하고, (2) 구축된 서지데이터의 품질을 객관적으로 분석하였으며, (3) 발견된 주요 문제점을 토대로 4가지 실행전략을 제안하였다. 이 연구에서 제안하는 실행전략은 서지데이터의 품질 제고 뿐만 아니라 업무효율과 데이터 공유를 위한 기반 구축을 동시에 실현하는 것을 목표로 하고 있다.

주제어 : 서지데이터, 서지레코드, 데이터 품질, 공공도서관, 부산

ABSTRACT : In 2020, the Busan metropolitan library took the lead in establishing an integrated library system (ILS) that integrates bibliographic data from 49 public libraries and 103 small public libraries. However, each library still builds bibliographic data individually and repeatedly, and the bibliographic data built by each library is only physically stored in an integrated DB. Therefore the improvement in work efficiency or data quality has not been achieved. This study aimed to analyze the construction processes and quality of bibliographic data in Busan public libraries and to suggest a new implementation strategy for an integrated environment. To this end, (1) the construction process of bibliographic data was investigated, (2) the quality of the constructed bibliographic data was objectively analyzed, and (3) four implementation strategies were suggested based on critical problems. The implementation strategy aims not only to improve the quality of bibliographic data, but also to increase work efficiency and build an infrastructure for data sharing.

KEYWORDS : Bibliographic Data, Bibliographic Records, Data Quality, Public Library, Busan

---

* 본 연구는『부산지역 도서관의 서지데이터 분석 및 서비스 활성화 방안 연구』의 일부를 수정・보완한 것임.
** 부산대학교 문헌정보학과 교수(jhrho@pusan.ac.kr / ISNI 0000 0004 6484 8385) (제1저자)
***동의대학교 문헌정보학과 조교수(ejulee@deu.ac.kr / ISNI 0000 0004 6335 8325) (교신저자)

• 논문접수: 2023년 8월 21일 • 최초심사: 2023년 8월 27일 • 게재확정: 2023년 9월 2일
• 한국도서관・정보학회지, 54(3), 105-128, 2023. http://dx.doi.org/10.16981/kliss.54.3.202309.105

※ Copyright © 2023 Korean Library and Information Science Society
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.

Ⅰ. 서 론

도서관 서비스의 기본은 자료에 대한 이용자의 효율적인 접근을 지원하는 데 있다. 이를 위한 일차적 도구가 목록(서지데이터)을 이용한 자료검색이라 할 수 있다. 방대한 장서 속에서 이용자가 원하는 자료를 탐색・식별하고, 관련된 다른 자료와의 관계를 파악할 수 있도록 일련의 데이터를 구조화하여 효율적인 검색서비스를 지원하는데 서지데이터의 궁극적인 목적이 있는 것이다. 아무리 중요하고 가치 있는 자료라 하더라도 이 과정이 제대로 되어 있지 않으면 자료의 존재조차 파악하기 어렵고, 자료가 가진 속성이 데이터로 적절히 표현되지 않으면 검색 자체가 불가능하다. 따라서 도서관에서는 이 업무를 전담하는 조직과 전문인력을 별도로 두고, 이를 도서관 서비스의 근간으로 삼고 있다.

오늘날의 서지환경에서 개별 도서관에서 구축한 서지데이터는 더이상 로컬에서의 사용만을 위해 존재하지 않으며, 나아가 로컬 데이터의 문제로만 그치지 않는다. 로컬 데이터는 지역과 국가, 국제적으로 상호연결된 거대한 서지네트워크에서 최하위 노드로서 기능하기 때문이다. 뿐만 아니라 시맨틱웹과 링크드 데이터 환경에 적합한 새로운 서지구조로의 전면적인 전환을 준비하고 있는 시점에, 현재의 서지데이터는 새로운 서지환경을 구축하는데 있어서도 성패를 결정하는 중요한 지표가 된다. 새로운 서지구조로의 전환을 위한 첫 번째 단계는 기존 데이터의 기계적 변환에서부터 시작하기 때문이다. 따라서 서지데이터의 품질은 당장의 활용뿐 아니라 차세대 서지구조의 실행을 위해서도 매우 중요한 문제라고 할 수 있다.

부산광역시 대표도서관인 부산도서관은 지역 도서관의 정책개발과 실행을 위한 광역대표도서관으로서 2020년에 개관하였다. 개관과 동시에, 국내 최초로 공공도서관의 운영주체나 규모에 상관없이 모든 도서관이 공동으로 사용할 통합도서관시스템(ILS)을 구축하고(부산도서관 누리집, 2023), 지역 내 49개 공공도서관과 103개 공립작은도서관의 서지데이터를 통합하였다. 그러나 3년이 경과한 현재까지도 각 도서관은 여전히 개별적으로 서지데이터를 구축하고 있고, 각 도서관에서 구축한 서지데이터는 물리적으로만 통합되어 있을 뿐 통합환경을 활용한 업무효율이나 데이터의 고품질화를 실행하지는 못하고 있다. 이러한 배경에서, 부산지역 공공도서관의 서지데이터 구축체계와 품질을 전반적으로 분석하고 통합환경에 적합한 새로운 구축방안을 모색할 필요가 있다.

이 연구는 2023년 현재 부산지역에 있는 49개 공공도서관을 연구범위로 한다. 실제 서지데이터는 공립작은도서관의 서지데이터까지 통합되어 있지만, 공립작은도서관의 경우 인력이나 운영상의 한계로 인해 데이터의 품질 면에서 공공도서관과 편차가 매우 큰 편이다. 따라서 공공도서관을 중심으로 서지데이터의 품질을 분석하고 실행방안을 마련하여, 이를 공립작은도서관까지 아우르는 방식으로 접근하는 것이 타당하다고 판단되어 공립작은도서관의 서지데이터에 대한 분석은 이번 연구에서 제외하였다. 분석에 필요한 데이터는 부산도서관으로부터 임시ID를 발급받아 부산

지역 도서관의 ILS에 직접 접속하여 수집하였다. 또한, 부산지역 공공도서관의 실무진 면담과 업무문서를 연구에 활용하였고, 문제점 도출 및 실행방안의 모색을 위해 국내외의 참조사례를 폭넓게 조사・분석하였다.

이 연구는 부산지역 공공도서관을 사례로 한 연구이지만, 이를 부산지역의 특수한 사례로만 한정하기 어렵다. 이번 연구를 진행하면서 국내 대부분의 공공도서관들이 부산지역 공공도서관과 유사한 구조로 업무를 수행하고 있고, 광역대표도서관을 중심으로 운영되는 통합DB도 대부분 부산시 사례와 크게 다르지 않은 것으로 확인되었기 때문이다. 이러한 점에서, 이 연구는 그동안 국가대표도서관이나 대학도서관을 중심으로 진행된 연구의 범위를 공공도서관의 서지데이터로 확장하고, 특히 열악한 공공도서관의 업무환경과 데이터의 문제에 대한 우리 학계와 도서관 현장의 관심을 촉발하고자 하는데 목적을 두고 있다.

Ⅱ. 서지데이터의 구축방식

부산지역 공공도서관의 경우 지역 기반의 '통합' ILS를 이용하여 KORMARC 형식으로 서지데이터를 구축하고 있다. ILS는 통합 플랫폼의 형태이지만 서지데이터는 도서관별로 각자 구축하고 있으며, 각 도서관에서 구축한 서지데이터는 하나의 DB에 통합 저장되는 구조이다. 이 시스템에서는 자관의 서지데이터만 이용할 수도 있고, 다른 도서관이나 부산지역 전체 도서관의 서지데이터를 실시간으로 자유롭게 검색하고 이용할 수도 있다.

현재 49개 공공도서관에서 서지데이터를 구축하는 방식은 다음 〈표 1〉과 같이 구분할 수 있다. 대표도서관인 부산도서관을 포함한 3개 도서관에서만 자관에서 직접 서지데이터를 구축하고 있을 뿐 대다수의 공공도서관에서 자료구입과 함께 MARC 형식의 서지데이터를 외부에서 반입(구입)하여 사용하고 있다. 서지데이터를 자체 구축하는 3개 도서관에서도 다국어 자료나 업무 상황에 따라 일부 자료에 대해서는 외부 데이터의 반입을 병행하기도 한다. 이에 반해, 통합 ILS를 이용하여 지역 내 다른 도서관의 서지데이터를 직접 카피(copy)하여 사용하거나 참조하는 경우는 없다. 주로 지역 도서관보다는 국립중앙도서관의 서지데이터를 직접 참조하기 때문이다. 따라서 통합 ILS는 지역 도서관 간에 서지데이터의 공유나 분담편목(shared cataloging)을 위한 용도로는 거의 활용되지 않는 상황이라 할 수 있다.

〈표 1〉 부산지역 공공도서관의 서지데이터 구축방식

직접 구축(original cataloging)	외부 데이터 반입	계
3개관 (6.1%)	46개관 (93.9%)	49개관 (100%)

외부 데이터를 반입하는 도서관에서는 기술적 요구사항이 수록된 '과업지시서'(또는 '도서구매 특수조건')를 데이터 구축 업체에 제공하고 이에 따라 서지데이터를 생성하도록 유도하고 있다. 과업지시서는 도서관마다 개별적으로 작성한 것이지만, 과업지시서에 담긴 도서관별 요구사항이나 조건에 큰 차이가 없고 대부분의 도서관들이 지역 내 동일 업체(서적상)로부터 데이터를 반입하고 있다. 따라서 각 도서관에서 구축한 데이터의 내용적 차이는 크게 없다고 볼 수 있다. 서지데이터의 구축 비용은 건(item)당 지불 방식이며, 복본이나 비도서자료(DVD자료, 전자책, 녹음자료 등)에 대해서도 동일한 비용을 지불하고 있다. 공공도서관의 경우 도서관 간에 자료의 중복 소장율이 매우 높고 한 도서관 내에서도 복본을 많이 소장하고 있다는 점을 고려할 때, 동일한 서지데이터를 반입하는데 상당한 비용을 중복지출하는 구조라 할 수 있다.

서지데이터는 모든 도서관에서 "1책당 1개의 레코드"를 원칙으로 구축되고 있다. 즉, 권차나 복본에 상관없이 모두 책 단위로 서지데이터를 구축한다. 이에 따라 통합 ILS에는 소장자료의 수만큼 서지데이터가 중복으로 저장된다. 예를 들어, A공공도서관에서 『공정하다는 착각』을 6권 소장하고 있다면 6건의 서지레코드가 구축되고, 동일한 방식으로 부산지역 전체 152개의 도서관(공립작은도서관 포함)에서 구축된 서지레코드가 중복적으로 통합DB에 저장되는 것이다. 이러한 서지데이터의 구축 원칙이 어떤 이유에서 언제부터 적용되었는지 추적하기는 어렵지만, 분명한 사실은 외부 데이터의 반입과 밀접한 관련이 있다는 점이다. 서지데이터의 반입이 '1책 1레코드'로 이루어지다보니 동일 자료라 하더라도 복본의 수만큼 서지데이터를 구축하는 현재의 구조를 갖추게 된 것이다.

문제는 자관에서 직접 서지데이터를 구축하는 도서관에서조차도 이러한 원칙을 동일하게 적용하고 있다는 데 있다. 이는 외부 데이터의 반입을 위한 고육지책으로 적용되던 방식이 자체 구축으로 회귀한 후에서도 여전히 유지되고 있음을 의미한다. 이에 대한 주요 이유를 조사한 결과, 실무진 면담에서 "이러한 구조에 대해 문제의식을 크게 가지지 않았고", "과거부터 이어져오던 원칙을 관행적으로 유지"하고 있기 때문으로 나타났다. 그러나 서지데이터의 중복 구축으로 인한 업무부담과 비효율성이 가중되고 있고,[1] 무엇보다 이렇게 구축된 서지데이터를 자료검색에 그대로 이용하다보니 검색의 효율 또한 크게 떨어지는 문제로 이어지고 있어 이에 대한 검토와 개선이 시급한 상황이라 할 수 있다(〈그림 1〉 참고).

한편, 부산지역 도서관들은 서지데이터의 구축을 위한 자체 지침(자료정리지침)을 가지고 있다. 서지데이터를 직접 구축하는 도서관에서는 이 지침을 근거로 데이터를 구축하며, 외부 데이터를 반입하는 도서관에서는 이를 과업지시서에 수록하여 외부 업체에 제공하고 있다. 따라서 지침을 분석하면 부산지역 도서관들이 구축한 서지데이터의 특징을 전반적으로 파악할 수 있다.

1) 〈그림 1〉에서 보듯이, 동일 자료에 대한 복수의 서지레코드를 비교해 보면 서지레코드의 구축 시점이나 업무담당자에 따라 데이터의 내용에 일부 차이가 있음을 확인할 수 있다. 이는 동일 자료라 하더라도 기존 서지레코드를 그대로 활용(copy)하는 것이 아니라 각각 새롭게 구축하고 있다는 의미로 해석할 수 있다. 이러한 상황은 외부 반입 데이터에서도 유사하게 나타난다.

〈그림 1〉 동일 자료에 대한 서지레코드 중복 구축 사례와 자료검색 결과 화면

49개 공공도서관 중에서 자체 구축 도서관 3개관과 외부 데이터 반입 도서관 4개관의 지침을 수집하여 분석한 결과, 구축방식에 따른 차이는 거의 없는 것으로 나타났으며, 서지데이터의 기술 수준은 공통적으로 KORMARC 기준으로 '최소수준'에 준하는 기본적인 데이터 중심으로 마련되어 있음을 확인할 수 있었다. 지침의 주요 내용은 접근점(표제부출의 생성, 저자 접근점과 이형의 기술 방법 등)과 여러 형태의 표제 기술방법, 주기사항의 기술 범위(8~9개의 필드 적용) 등으로 구성되어 있었고, 각 항목의 내용은 전체적으로 매우 간소한 편이었다. 또한, 활용도가 높은 데이터임에도 불구하고 부호화 정보(리더, 006, 007, 008 필드)에 대한 내용은 대부분 누락되어 있었고, 심지어 일부 지침에서는 오류가 발견되기도 하였다. 이처럼 서지데이터 입력지침은 데이터의 일관성을 유지하고 업무효율을 증대하기 위한 목적에서 부산지역 공공도서관에서 보편적으로 활용되고 있었지만, 역설적이게도 지침이 가진 내용적 빈약성과 오류가 서지데이터의 품질 문제를 야기하는 원인으로도 작용할 수 있는 상황임을 확인할 수 있었다. 서지데이터의 구축방식이나 업무구조 등에 대한 전반적인 현황 파악에 이어, 다음 장에서는 실제 서지데이터의 품질을 세밀하게 들여다보고자 한다.

Ⅲ. 서지데이터의 품질

1. 분석방법

데이터의 품질은 곧 데이터의 집합인 데이터베이스의 품질 측면에서 살펴볼 필요가 있다. 다양한 유형의 데이터베이스 품질을 평가하기 위해 국내외의 많은 연구자나 연구기관에서 평가모델을 개발하여 적용해 왔는데, 그러한 평가모델은 모든 상황에 공통으로 적용할 수 있는 단일의 모델이 아니라 품질평가의 목적이나 주체, 결과의 활용도 등에 따라 차이를 보인다. 국내에서는 대표적으로

한국데이터베이스진흥센터에서 국내외 관련 연구를 추적하여 데이터베이스의 품질을 평가할 수 있는 객관적이며 계량적인 평가방법을 종합하여 소개한 바 있다(한국데이터베이스진흥센터, 2000). 한국데이터베이스진흥센터에서는 데이터베이스의 품질을 "데이터의 품질 + 서비스의 품질 + 기타 품질"로 구분하고 품질분석을 위한 세부 기준을 제시하였는데, 데이터의 품질기준은 주로 '데이터 자체의 정확성', '논리적/물리적 완전성', '최신성', '일관성' 등이 사용되는 것으로 나타났다.

데이터가 가진 정확성이나 무결성 등이 초기의 품질평가 방법론에서 절대적으로 중요한 기준이었다면, 비교적 최근에는 실제 데이터의 활용 측면, 즉 데이터가 사용자의 요구를 만족시키는지에 대한 중요성을 부각하는 경향을 보이고 있다. 따라서 데이터의 '유용성', '적시성', 데이터에 대한 '사용자 만족도' 등이 데이터의 품질기준에 새롭게 포함되기도 하였다(한국정보화진흥원, 2018). 국내 도서관의 서지데이터 또는 종합목록을 대상으로 품질평가를 시도한 선행연구들도 대부분 이러한 범주에서 크게 벗어나지 않는 것으로 확인되며(김선애, 이수상, 2006; 박진호, 2018; 양수완, 2020; 이제환, 2002), 국외 연구 중에는 설문조사나 인터뷰 등을 통해 분석목적에 적합한 품질기준을 별도로 선정하거나 이용자, 사서, 업무조직 등 다양한 관점을 품질기준에 반영하고자 한 시도도 찾아볼 수 있다(Kyprianos, Lolou, & Efthymiou, 2022; Snow, 2011).

이 연구에서는 이러한 국내외 선행연구를 참고하여 부산지역 공공도서관에서 생산한 서지데이터의 품질을 분석하기 위한 기준으로 "표준이나 지침을 준수하여 얼마나 정확하고 완전하게 작성되었는지", 그리고 "동일한 데이터를 도서관 간에 또는 자료별로 얼마나 상이하게 표현하고 있는지"로 설정하였다.[2] 분석할 세부항목은 부산지역 도서관의 통합환경을 고려하여 자료검색에서 중요하게 또는 직접적으로 사용되는 데이터(예를 들어, 표제정보나 저자정보, 내용식별을 위한 자원유형이나 형태정보, 발행년도 등)나 향후 새로운 서지구조로의 전환에서 특별히 중요하게 사용될 데이터(원표제, 언어정보 등)를 중심으로 선정하였다. 분석결과는 데이터 자체의 절대적인 품질을 분석하는데 중점을 두되, 분석기준에 따라 정량적 수치(레코드 건수나 비율) 또는 관련 사례를 제시하는 것을 목표로 하였다. 최종적으로, 이 연구에서 서지데이터를 분석하기 위해 설정한 기준과 세부항목은 다음 〈표 2〉와 같다.

〈표 2〉 서지데이터 분석기준과 세부항목

분석기준	내용	세부항목
정확성	오류나 누락 없이 데이터가 정확하게 입력되었는가?	• 데이터의 오류 • 데이터의 누락 • 데이터의 불일치

2) 데이터의 품질기준에 주로 포함되는 '갱신주기'나 '신속성'(최신성 혹은 적시성)은 서지데이터만으로는 정확한 분석이 불가능하기 때문에 이번 분석에서는 제외하였다. '중복성' 또한 앞서 언급하였듯이 권책 단위의 서지레코드 구축 방식에서는 의미가 없다고 판단하여 분석기준에서 제외하였다.

분석기준	내용	세부항목
완전성	자원의 속성이 완전하게 표현되었는가?	• 자원유형이나 형태 정보 • 내용 설명정보 • 주제적 속성
표준의 준수	표준이나 내부 지침을 준수하고 있는가?	• 표준의 미준수 • 자체 지침의 미준수
기술방식의 차이	동일한 데이터를 표현하는 방식에 차이가 있는가?	• 표제정보의 기술방식 • 저자정보의 기술방식

2. 샘플 데이터의 선정

부산지역 공공도서관에서 구축한 서지데이터의 품질을 분석해보기 위해 이 연구에서는 샘플 도서관을 먼저 선정한 후 해당 도서관의 서지데이터를 추출하는 과정을 거쳤다. 부산지역 공공도서관은 서지데이터의 구축방식에 따라 직접 구축하는 도서관(3개)과 MARC 반입하여 사용하는 도서관(46개)으로 구분할 수 있으며, 설립주체에 따라 지자체 도서관(35개)과 교육청 소속 도서관(14개)으로 구분할 수 있다.[3] 이 연구는 도서관 간의 상대적 비교가 아니라 전체적인 품질 수준을 가늠하는데 목적이 있으므로 서지데이터의 구축방식과 설립주체를 기준으로 도서관을 1차 구분하고, 그 중에서 소장자료의 규모가 상대적으로 큰 도서관을 분석대상으로 선정하였다. 분석대상 도서관의 현황은 다음 〈표 3〉과 같다.

〈표 3〉 서지데이터 분석대상 도서관의 선정

서지데이터 구축 방식	설립주체		분석대상 도서관
(1) 직접 구축: 2개관	(1) 지자체 공공도서관: 4개관	⇨	**총 6개관**
(2) 외부 반입: 4개관	(2) 교육청 소속 공공도서관: 2개관		

분석대상 도서관의 서지데이터는 통합 ILS에서 구분하고 있는 자료구분, 즉 '단행/연속자료'와 '비도서자료'로 나누어 추출하였다. 단행/연속자료에는 단행본, 큰활자도서, 점자자료, 연속간행물이, 비도서자료에는 전자책, 비디오녹화자료(DVD), 녹음자료, 웹DB, 컴퓨터파일 등이 포함된다. 자료유형별로 샘플 레코드를 추출한 기준은 다음과 같다.

먼저, 단행/연속자료에서는 각 도서관에서 공통으로 소장한 도서(단행본, 점자자료, 큰활자자료) 중 비교적 최신의 자료를 선정하고자 하였다. 이에 「도서관 정보나루」에서 최근 10년간(2013-2022년)

3) 공공도서관의 설립주체에 따른 데이터 품질의 차이를 분석하기 위한 목적에서 설립주체별로 분석대상 도서관을 선정하였으나, 결과적으로 설립주체에 따른 유의미한 차이는 없는 것으로 나타났다. 따라서 연구결과에는 서지데이터의 구축 방식에 따른 차이만 서술하였다.

부산지역 20대 이상의 인기대출도서 중 철학, 사회과학, 문학, 역사 분야의 자료를 선정하고, 이 가운데 다양한 표제를 가지거나 원자료의 언어가 상이한 자료 등 12종을 임의로 선별하였다. 분석대상 도서관인 6개 도서관에서 12종의 자료에 대한 서지데이터는 총 336건으로 나타났다(자체 구축 97건, 외부 반입 239건). 다음으로, 비도서자료는 도서관들마다 소장하는 자료의 유형이 다양하고, 이에 따라 공통으로 소장하고 있는 자료로 선정하기가 어려워 부득이 임의의 검색어('한국사회' 및 '사상')를 이용하여 추출하는 방식을 사용하였다. 분석대상으로 선정한 6개 도서관에서 추출한 비도서자료의 서지레코드는 총 189건이었다(자체 구축 60건, 외부 반입 129건). 샘플 레코드의 전체 현황은 〈표 4〉와 같다.

〈표 4〉 샘플 레코드의 선정

	직접 구축		외부 반입				계
	A도서관	B도서관	C도서관	D도서관	E도서관	F도서관	
도서	57	40	47	28	92	72	336
소계	97		239				
비도서	43	17	10	14	81	24	189
소계	60		129				

3. 품질분석 결과

가. 정확성

첫 번째 분석기준인 데이터의 정확성은 ① 데이터의 오류, ② 데이터의 누락, ③ 데이터의 불일치로 구분하여 분석하였다.

먼저, 데이터의 오류는 단순 철자오류나 정보원에 있는 정보를 잘못 입력한 표기오류, 기타 필드 사용 오류 등을 오류로 간주하되, 띄어쓰기 오류나 수량의 미세한 차이(예: 도서의 페이지)는 분석에서 제외하였다. 결과적으로, 도서자료의 서지레코드에서 발견된 오류 중 가장 빈번하게 발견된 데이터의 오류 유형은 '발행년도' 오류로 나타났다. 동일 자료의 동일 판에 대한 발행년도를 제각각 다르게 기재한 사례가 가장 많았는데, 이는 아래와 같이 발행년도가 아닌 재쇄년도나 인쇄년도 등을 임의로 입력한 명백한 오류에 해당하는 사례였다.

[동일 자료에 대한 상이한 발행년도 입력 사례]

260 ▼a서울: ▼b문학사상사, ▼c2009

260 ▼a서울: ▼b문학사상사, ▼c2011

260 ▼a서울: ▼b문학사상사, ▼c2012

260 ▼a서울: ▼b문학사상사, ▼c2010 ▼g2012(2판 33쇄)

비도서자료의 서지레코드에서는 자료유형표시(GMD)와 특정자료종별(SMD)에서 특히 많은 오류가 발견되었다(〈표 5〉 참고). 비도서자료의 경우 동일한 표제를 가진 상이한 매체를 명확하게 구분하기 위해 GMD와 SMD의 표기 형식을 엄격히 적용하고 있고, 나아가 새로운 서지구조에서도 이들 데이터를 개체 구분을 위한 중요한 단서로 활용하고 있음을 고려할 때 이는 결정적인 오류에 해당한다고 볼 수 있다.

〈표 5〉 전자책의 GMD와 SMD 데이터의 오류

올바른 기술 사례	오류 사례
245 00▼a청춘 심리 상담▼h**[전자자료]** : ▼b병든 한국 사회는 청년들의 내면을 어떻게 파괴했는가 / ▼d김태형 지음 256 ▼a**전자 데이터** 260 ▼a서울 : ▼b다시봄, ▼c2016 ▼e(서울 : ▼fYES24, ▼g2016) 300 ▼a**전자책 1책**(263 p.) : ▼b천연색 삽화	245 ▼h**[전자책]** 380 ▼a**전자책** 380 ▼a**오디오북** 300 ▼a**전자자료**(Application) : ▼b삽화; ▼c**전자책 1책** 300 ▼ae-book(5users) : ▼bXML 300 ▼a34M(E-book, 1user) : ▼b천연색 300 ▼aE-BOOK(5User) : ▼bXML

도서자료와 비도서자료에서 발견된 오류 데이터의 전체 현황은 다음 〈표 6〉과 같다.

〈표 6〉 오류 데이터의 비율

구분	유형		오류 데이터의 비율	구축주체	
				직접	반입
도서	철자오류		1.5%	0.6%	0.9%
	표기오류	발행년도	25.0%	35.1%	20.9%
		발행지	1.8%	0.0%	12.0%
비도서	표기오류	자료유형표시	65.6%	89.0%	56.6%
		특정자료종별과 수량	69.8%	90.0%	60.5%
		발행지	0.5%	0.0%	14.9%
	필드사용 오류	저작의 형태(380필드)	18.0%	100%	0.0%

다음으로, 데이터의 누락은 도서자료의 경우 번역본에 대한 '원표제'의 누락(17%)과 '언어정보'(041 필드)의 누락(18.9%)이 가장 빈번하게 나타났다. 비도서자료에서는 자료유형표시의 누락(11.1%), 발행지나 배포지의 누락(2.1%) 등이 발견되었다.

마지막으로, 데이터의 불일치는 텍스트 데이터와 부호 값의 불일치를 의미하며, 샘플 레코드에서 가장 빈번하게 발견된 데이터의 불일치 유형은 008/15-17(발행국명) 부호와 260 필드의 ▼a(발행지, 배포지 등)의 불일치였다. 발행지 정보를 텍스트로는 '파주', 부호 값으로는 'bnk'(부산)으로 입력한 사례가 이에 해당되었다. 도서자료에서 8.6%(총 336건 중 24건)가, 비도서자료에서 단 1건

이 이에 해당하는 사례였다.

데이터의 정확성을 데이터의 구축방식에 따라 비교한 결과, 전체적으로 직접 구축 데이터가 외부 반입 데이터에 비해 정확성이 약간 높은 편이었지만, 전체적으로 유의미한 통계적 차이를 보이지는 않았다. 가령, 〈표 7〉과 같이 도서자료에서 언어정보(041 필드)의 누락과 비도서자료에서 자료유형표시(245필드의 ▼h)의 누락을 데이터의 구축 방식에 따라 비교한 결과를 보더라도, 직접 구축 데이터와 반입 데이터 간 평균의 차이는 크지 않다. 오히려 직접 구축 도서관들 간에 혹은 데이터 반입 도서관들 간에 편차가 크다는 것을 알 수 있다. 특히, 서지데이터를 직접 구축하는 두 도서관을 비교해 보면, A도서관 데이터의 정확성이 모든 면에서 전반적으로 매우 높게 나타난 반면 B도서관의 경우 6개 도서관 중에서 가장 낮다. 이러한 결과를 토대로 볼 때, 구축방식에 따른 품질의 절대적 우위를 단정하기는 어렵고, 이보다는 실무자의 개인적 역량이나 외부로부터 반입한 데이터에 대한 엄격한 검수 여부가 서지데이터의 품질에 더 큰 영향을 미치는 것으로 예측할 수 있다.

〈표 7〉 데이터 구축 방식에 따른 데이터의 누락율 비교

구분		직접 구축		데이터 반입				계
		A	B	C	D	E	F	
도서	해당 레코드의 수(건)	36	27	35	21	37	50	156
	041 필드의 누락(건)	0	11	5	3	6	14	66
	041 필드의 누락 비율(%)	0	40.7	14.3	14.3	16.2	28.0	18.9
	구축 방식에 따른 평균(%) 비교	17.5		19.6				-
비도서	해당 레코드의 수(건)	43	17	10	14	81	24	189
	자료유형표시(245 ▼h) 누락(건)	0	6	4	0	8	3	21
	자료유형표시(245 ▼h) 누락 비율(%)	0.0	35.3	40.0	0.0	9.9	12.5	11.1
	구축 방식에 따른 평균(%) 비교	10.0		11.6				-

나. 완전성

데이터의 완전성은 ① 자원의 유형이나 형태 정보, ② 내용 설명정보, ③ 주제적 속성 측면에서 데이터가 얼마나 상세하고 적절한지를 중심으로 살펴보았다.

먼저, 자원의 유형이나 형태에 관한 정보는 자료의 형태나 물리적 특성에 관한 상세한 정보를 입력하는 KORMARC의 007(형태기술정보) 필드에 입력된 데이터를 분석하였다. 특히, 007 필드는 자원의 범주를 식별하거나(007/00 - 자료범주표시) 해당 범주에 속하는 자료의 세부적인 유형(007/01 - 특정자료종별)을 형식화된 부호로 입력하기 때문에 검색에 활용하기에 매우 용이한 구조로 되어 있다. 일반도서의 경우 007 필드에 입력된 데이터는 크게 의미가 없다고 볼 수도 있으나 이 데이터를 통해 일반도서, 큰활자도서, 점자자료 등을 구분할 수 있고, 다양한 매체로 된 자원은 각 자원의 유형과 물리적 특성을 더욱 분명하게 식별할 수 있어 통합검색 환경에서

이들 데이터의 활용도는 매우 높다고 할 수 있다. 이러한 이유로 부산지역 일부 도서관의 입력지침에서도 영상자료, 전자자료, 녹음자료 등에 대해서는 007 필드의 적용을 '필수'로 설정하고 있다.

샘플 데이터에서 007 필드의 사용빈도를 확인한 결과, 도서자료는 336건 중 절반가량인 170건(50.6%)에만 적용되어 있었고, 비도서자료는 189건 중 57건(30.2%)에 불과해 매체 구분이 더 필요한 비도서자료에서 관련 데이터의 기술이 더욱 불완전한 것으로 나타났다(〈표 8〉 참조). 007 필드가 입력된 서지데이터의 비율은 외부 반입 데이터보다 도서관에서 직접 구축한 데이터가, 그 중에서도 특히 A도서관의 데이터에 비교적 완전하게 기술되어 있었다.

〈표 8〉 007(형태기술필드) 필드의 사용

구분		직접 구축		외부 반입				계
		A	B	C	D	E	F	
도서	007 필드의 사용 비율(%)	87.7	22.5	58.8	28.4	43.5	51.4	50.6
	구축 방식에 따른 비율(%) 비교	60.8		44.8				
비도서	007 필드의 사용 비율(%)	81.4	11.8	0.0	0.0	6.2	62.5	30.2
	구축 방식에 따른 비율(%) 비교	61.7		15.5				

다음으로, 자료의 내용을 설명하는 정보가 얼마나 자세하게 기술되었는지를 살펴보기 위해 5XX 필드(내용주기, 이용대상자주기, 수상주기 등)를 분석하였다. 5XX 필드에는 자료의 성격이나 범위, 자료의 특성을 설명하거나, 표제와 책임표시사항, 판사항, 발행사항, 형태사항, 총서사항에 기술하기에 적절하지 않거나 여기에 기술된 내용을 설명・보완・한정하는 데이터를 기록하고, 이를 자원에 대한 식별정보로 활용하고 있다. KORMARC이나 KCR4에서는 주기사항에 기록될 수 있는 데이터를 유형별로 구분하고 있는데, 주기사항에 기록될 데이터의 유형은 매우 많은 편이다(KORMARC 기준으로 총 51개 필드).

샘플 레코드에서 단 한 번이라도 사용된 주기사항의 사례를 분석한 결과, 도서자료에서 총 11개 필드가, 비도서자료에서 총 15개 필드가 사용된 것으로 나타났다. 도서자료에서는 500(일반주기), 504(서지 등 주기), 521(이용대상자주기), 546(언어주기) 필드의 사용빈도가 높은 편이었고, 부산지역 도서관들의 입력지침에 명시된 주기사항 필드 중 505(내용주기), 506(이용제한주기), 520(요약 등 주기) 필드 등은 단 한 건도 사용되지 않은 것으로 나타났다. 비도서자료에서는 511(연주자와 배역진 주기), 520(요약 등 주기), 521(이용대상자 주기), 538(시스템사항에 관한 주기) 필드가 모든 도서관에서 공통적으로 사용되고 있었다.

주기사항에 입력된 데이터로 보았을 때 자료에 대한 충분한 정보를 제공하고 있다고 보기는 어려웠고, 특히 아래와 같이 주기사항이 아니라 다른 항목(발행사항, 판사항 등)에 기재해야 하는 데이터이거나 "한학기한권읽기", "2022년 △△도서관에서 이관" 등과 같이 자원 자체와 관련 없는 정보가 주기사항 필드에 기술되어 있는 사례도 다수 확인되었다.

[500(일반주기) 필드에 기술된 데이터의 사례]
기시미 이치로의 한자명은 "岸見一郎"이고, 고가 후미타케의 한자명은 "古賀史健"임
특별판
이 책은 점자도서임
점자본제작: 한국점자도서관(2018년 03월 제작)
등록번호 EM6776은 2006년에 발행됨
2019년 세종도서 교양부문 선정도서임
2022년 △△도서관에서 이관
한학기한권읽기

마지막으로, 자료의 주제적 속성을 살펴본 결과, 주제명은 도서와 비도서 모두에서 통제어휘보다 비통제어휘(653 필드)로 작성된 데이터가 월등히 많은 것으로 나타났다(〈표 9〉 참조). 비통제어휘로 입력된 데이터는 〈표 10〉과 같이 전체적으로 표제에서 직접 추출한 단어의 비중이 높았다. 도서자료의 경우, '어떻게', '얻는가', '도망친' 등과 같이 주제명으로 사용하기에 부적합한 단어들이 다수 포함되어 있었고, 비도서자료에서는 '전자책', '이북', 'ebook', '오디오북', '화면해설영상', '전자오디오북', '전자자료' 등 자료유형이나 자료의 성격을 나타내는 단어들이 상위 주제명으로 나타났다. 통제어휘로 입력된 주제명 데이터만 별도로 확인한 결과, 통제어휘로 된 주제명은 모두 국립중앙도서관의 데이터를 그대로 사용(copy)한 것으로 확인되었다.

〈표 9〉 주제명 데이터의 유형

주제명 데이터의 유형	도서	비도서
통제어휘로만 작성된 레코드 비율(%)	2.4	0.5
비통제어휘로만 작성된 비율(%)	**82.4**	**90.5**
통제어휘와 비통제어휘가 동시에 존재하는 비율(%)	13.1	9.0

〈표 10〉 비통제 주제명 데이터의 사례

샘플 자료의 표제	주제명과 사용빈도			
창문 넘어 도망친 100세 노인: 요나스 요나손 장편소설	**창문**	36	**도망**	17
	100세	34	**도망친**	17
	노인	34	스웨덴문학	5
	장편소설	23	**100세노인**	2
	요나손	18	모험	2
	요나스	18	스웨덴소설	2
총, 균, 쇠: 무기·병균·금속은 인류의 운명을 어떻게 바꿨는가	**인류**	23	**쇠**	11
	금속	22	**총**	11
	무기	22	**인류문명**	7
	병균	22	문명사	2
	운명	18	진화론	2
	총균쇠	15	문명	1
	문명생성	14	제러드다이아몬드	1
	균	11		

다. 표준의 준수

표준의 준수에서는 ① KORMARC 표준과 ② 자체 입력지침의 준수로 나누어 살펴보았다.

먼저, KORMARC(2014 개정)에서 사용이 중단된 필드인 1XX(기본표목) 필드와 440(총서사항/부출표목) 필드의 사용은 다음 〈표 11〉과 같이 나타났다. 이들 필드는 부산지역 도서관들의 자체 입력지침이나 과업지시서에서도 사용하지 않는다고 분명하게 명시되어 있다. 그러나 사용 비율이 높은 편이었고, 특히 440 필드의 사용 비율이 도서자료와 비도서자료 모두에서 상당히 높게 나타났다.

〈표 11〉 KORMARC 표준의 미준수

구분	도서	비도서
1XX 필드를 사용한 레코드의 비율(%)	26.2	34.9
440 필드를 사용한 레코드의 비율(%)	88.9	41.4

한편, 부산지역 공공도서관들의 자체 지침에서는 표제 검색을 위한 접근점의 생성을 위해, 245 필드의 ▼b(표제관련정보), ▼x(대등표제), ▼p(권차표제)를 모두 740 필드에 부출하도록 규정하고 있다. 이에 대한 실제 적용 비율을 분석한 결과, 245 필드의 ▼b는 대부분 740 필드를 이용하여 부출하고 있지만, 245 필드의 ▼x나 ▼p를 부출한 사례는 비교적 적은 것으로 나타나 지침과는 다소 상이하게 적용되고 있음을 확인할 수 있었다. 지침의 준수여부와 상관없이, 지침에 명시된 바에 따라 다양한 표제를 모두 부출(분출)을 하는 것이 적절한지에 대해서는 추가적인 검토가 필요한 상황으로 보인다.

245 10▼a어떻게 원하는 것을 얻는가: ▼b**13년 연속 와튼스쿨 최고 인기 강의**/▼d스튜어트 다이아몬드 지음; ▼e김태훈 옮김
740 2▼a**13년 연속 와튼스쿨 최고 인기 강의**

245 10▼a왓칭 = ▼xWatching : ▼b신이 부리는 요술 / ▼d김상운 지음
740 ▼a**신이 부리는 요술**
740 ▼a**Watching**

245 10▼a영상실록, 20세기 거장들▼h[DVD녹화자료].▼n[2-1]/▼d다우리엔터테인먼트 [제작]
246 13▼a문학과 사상을 꽃피운 사람들(Literature and Thought)
740 2▼a**문학과 사상을 꽃피운 사람들(Literature and Thought)**

라. 기술방식의 차이

기술방식의 차이는 데이터의 오류에 해당하지는 않지만, 도서관별로 동일한 데이터를 얼마나, 어떻게 상이하게 적용하고 있는지를 분석하는데 목적이 있다. 이는 부산지역 도서관과 같은 통합환경에서 분석의 필요성이 대두되는 부분이라 할 수 있다.

분석의 대상을 '표제'와 '저자'로 한정하여 살펴본 결과, 표제 중에서는 표제관련정보나 대등표제의 기술방식에서, 저자명에서는 외국인명의 표기방식에서 도서관별로 많은 차이가 있었고, 하나의 도서관 내에서도 동일 자료에 대한 복수의 서지레코드에 상이하게 적용된 사례도 적지않게 발견되었다.

[상이한 표제 기술 사례]

사례 ① 245 10▼a왓칭:▼b신이 부리는 요술/▼d김상운 지음

사례 ② 245 10▼a왓칭:▼b신이 부리는 요술/▼d김상운 지음
246 1 ▼a**Watching**

사례 ③ 245 10▼a왓칭=▼xWatching:▼b신이 부리는 요술/▼d김상운 지음
246 13▼a**신이 부리는 요술**

저자명에 대해서도 하나의 접근점과 복수의 이형으로 구분하는 경우(예: 700 필드와 900필드)나, 상이한 이름 표기 형식을 모두 동일한 접근점 필드(예: 700 필드)에 반복 적용한 경우도 있었다. 또한, 외국인명의 경우 원어음의 한글표기, 한자의 한글음 표기, 원어 표기 등이 도서관마다 또는 한 도서관 내에서도 서지레코드마다 상이하게 적용되어 있는 것으로 나타났다. 일반적으로 저자명은 저자 접근점(7XX)에 전거형으로 제어된 형식을 기술하고, 다양한 형식의 이름을 전거레코드를 통해 제어한 후 이를 검색에 활용하는 방식을 적용한다. 그러나 전거데이터를 구축하지도, 외부의 전거데이터를 사용하지 않는 부산지역 도서관에서는 서지레코드에 저자명의 다양한 형식을 이처럼 일관성 없이 적용하는 것으로 나타났다.

[저자명의 상이한 적용 사례]

사례 ① 700 1 ▼a**기시미 이치로**
900 11▼a안견일랑

사례 ② 700 1 ▼a**岸見一郞**
700 1 ▼a**기시미 이치로**

사례 ③ 700 1 ▼a**기시미 이치로**▼d**1956-**
900 10▼aKishimi, Ichiro▼d1956-
900 10▼a안견일랑▼d1956-

Ⅳ. 서지데이터의 문제점과 품질 제고 방안

1. 문제점

부산지역 공공도서관은 도서관의 규모나 운영주체에 상관없이 하나의 도서관시스템으로 연결되어 있기 때문에 필요한 경우 타 도서관의 데이터를 참고하고 활용할 수 있는 체제를 갖추고 있다. 그러나 도서관들 간에 서지데이터의 실질적인 공유나 분담편목은 전혀 이루어지지 않고 있고, 152개의 공공도서관과 공립작은도서관에서 모두 개별적으로 서지데이터를 구축하는 방식을 여전히 고수하고 있다. 이에 따라 애초에 목표한 업무효율이나 통합검색 환경은 제대로 실행되지 못하고 있다. 부산지역 공공도서관 49개관 중에서는 3개관을 제외한 도서관에서 자료구입과 함께 서지데이터를 외부로부터 반입하여 사용하고 있는데, 서지데이터의 품질을 비교한 결과 유사한 수준의 서지데이터를 반입하기 위해 도서관마다 각각 비용을 지출하고 있어 부산지역 도서관 전체적으로 적지 않은 비용이 중복투자되고 있음도 이번 연구에서 확인할 수 있었다.

서지데이터의 구축방식에 따른 품질의 차이를 살펴본 결과, 이번 연구의 샘플링이 갖는 한계로 일반화하기는 어렵지만 도서관 내에서 서지데이터를 직접 구축하는 도서관, 그 중에서도 A도서관의 품질이 전반적으로 우수한 것으로 나타났다. 서지데이터를 직접 구축하는데 투입되는 인력과 시간이 MARC 반입과 비교하여 체감하는 업무량이 “15-20% 정도 증가한 수준”이라는 실무진의 의견을 참고할 때 업무부담이 감당하지 못할 만큼 높은 것은 아닌 것으로 나타나,[4] 서지데이터의 반입이 데이터의 품질이나 투입비용 측면에서 효과적이라 보기는 어려운 수준으로 나타났다. 따라서 서지데이터의 품질을 향상하기 위한 방책으로 도서관 실무진이 서지데이터의 구축에 직접적으로 관여하고 그 결과를 공유함으로써 업무효율을 증대하기 위한 방안을 지역 차원에서 공동으로 모색하는 것이 타당하다고 판단된다.

서지데이터의 공유와 이를 통한 예산과 인력, 업무부담의 경감을 위해서는 일반적으로 두 가지 환경적 조건이 충족되어야 한다. 첫 번째는 서지데이터의 공유를 위한 확실한 거버넌스의 구축이고, 두 번째는 효율적인 데이터의 공유나 분담편목을 지원하는 서지유틸리티의 기능 구비라 할 수 있다. 부산지역의 경우 현재 서지데이터의 공유나 분담편목을 위한 거버넌스가 존재하지 않는다. 일반적으로 종합목록을 운영하기 위해서는 제도와 시스템, 데이터 등을 통합관리하는 전담부서나 인력을 갖추고, 참여도서관으로 구성된 실무진이 참여하여 관련 정책을 개발하고 집행하는 체제를 갖추어야 한다. 그러나 부산지역 도서관은 시스템적으로 데이터의 통합만 이루었을 뿐 대표도서관을 중심으로 이러한 역할을 수행하는 실질적인 조직이나 제도를 운용하지 못하고 있다.

---

4) A도서관에서는 4명의 인력이 서지데이터 구축을 담당하고 있으며, 2022년 기준 연간 도서증가량이 50,095책(구입 32,512책, 기증 17,583책)이다(국가도서관통계시스템, 2023).

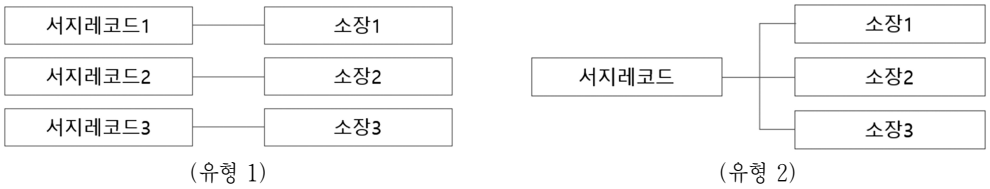
또한, 통합 ILS를 통해 각 도서관의 서지데이터는 시스템적으로 상호 연계되어 있기 때문에 타 도서관의 데이터를 실시간 활용할 수 있지만, 이를 실질적인 서지유틸리티로 보기는 어려운 상황이다. 서지유틸리티의 기본적인 기능은 서지네트워크에 참여하는 도서관에서 구축한 서지DB를 통합 관리하면서 이를 필요로 하는 도서관들이 서지데이터 및 관련 서비스에 실시간으로 접근할 수 있도록 편의 기능을 제공하는데 있다. 서지유틸리티의 세부기능은 운영주체에 따라 다소 차이가 있지만, 일반적으로 고품질의 서지레코드(master records, 대표서지)를 유지하면서 신규 서지레코드의 입력과 기존 서지레코드의 수정 및 편집, 다양한 형식의 목록 제작, 전거파일의 공동 활용, 관련 서비스의 제공 등의 기능을 제공하는 방식을 취하고 있다. 그러나 부산지역에서는 이러한 기능을 수행하는 서지유틸리티가 현재 마련되어 있지 못한 상태이다.

다만, 부산지역 도서관과 같은 업무구조에서, 지역 차원에서 서지유틸리티를 운영하는 것이 과연 비용효과적이며, 나아가 바람직한가에 대해 진지하게 고민해 볼 필요가 있다. 서지유틸리티를 운영하기 위해 적지 않은 인력과 비용이 투입되어야 하고, 또 과거에 지역 단위에서 긴밀한 협조 하에 운영되던 서지유틸리티가 DB의 '포괄성' 문제로 인해 결국 더 방대한 서지유틸리티로 통합된 국외 도서관의 사례를 참조해 보더라도 지역 차원에서 운영하는 서지유틸리티가 반드시 비용효과적이라 예측할 수는 없기 때문이다. 더구나, 부산지역 도서관들이 지역 기반 서지데이터의 공유에 대한 의지와 요구가 높지 않은 상황이고, 현재도 공공도서관의 분담편목을 지원할 목적으로 구축된 KOLIS-NET보다 국립중앙도서관의 데이터를 직접 참조하는 경향이 월등히 높게 나타나고 있어 지역에서 추진하는 서지유틸리티의 구축 및 운영이 지역 도서관의 업무부담을 경감하고 서지레코드의 품질을 제고하기 위한 해결책이 될 수는 없다고 판단되기 때문이다.

한편, 부산지역 도서관에서 구축한 서지데이터의 가장 근본적인 문제는 데이터의 구조로 파악되었다. 도서관에서 서지데이터를 직접 구축하거나 아니면 외부 데이터를 반입하거나 하는 방식에 상관없이, 모든 도서관에서 책 단위로 서지데이터를 구축하는 방식은 업무효율 및 DB 관리를 크게 저해할 뿐만 아니라 앞서도 언급하였듯이 외부 데이터를 반입하는 경우 동일한 서지데이터에 대해 비용을 중복으로 지출하는 주요 원인이 된다. 나아가, 이러한 구조로 된 서지데이터가 부산지역 도서관의 통합DB에 별도의 처리 없이 그대로 축적되다보니 통합DB의 품질은 물론이고 기본적인 검색과정에서도 많은 문제를 일으키고 있다.

부산지역 도서관들이 채택하고 있는 이러한 서지데이터의 구조를 국내외 공공도서관의 사례와 비교해보면 다음과 같다. 우선, 국내의 대표적인 사례로 광역대표도서관을 조사한 결과, 부산지역 공공도서관들과 마찬가지로 동일 자료(복본)를 책 단위로 분리 구축하는 방식(유형 1)이 절대적으로 많은 것으로 나타났다. 구체적으로, 부산지역을 제외한 16개 광역대표도서관 중에서 서울도서관, 대전한밭도서관 등 2개관만이 하나의 서지레코드에 복수의 소장정보를 연결하는 통합구축

방식(유형 2)을 적용하고 있을 뿐이다(〈그림 2〉 참고). 통합구조로 작성된 서지데이터는 검색에서도 하나의 서지레코드에 복수의 소장정보가 연결된 형태로 구현되지만, 분리 구축된 서지데이터는 각 자료별로 서지데이터가 제공되기 때문에 중복과 과다한 검색결과로 인한 식별력이 매우 떨어진다는 한계를 보이고 있다.



〈그림 2〉 동일자료에 대한 서지레코드의 구조

북미권 주요 도시의 사례 또한 '유형 2'의 구조로 되어 있음을 확인할 수 있다. 특히 북미권 공공도서관은 1개의 중앙도서관(central library)과 다수의 분관(branches) 체제로 운영되고 있고, 중앙도서관에서 분관의 자료를 포함하여 모든 자료에 대한 서지데이터를 통합 구축하여 중앙도서관의 자료검색시스템에서 모든 도서관의 자료를 통합 검색하는 구조를 갖고 있다. 따라서 하나의 서지레코드에 다수의 소장정보(소장도서관, 소장위치, 이용정보 등)를 연결하여 관리하되, 다수의 분관에 소장된 소장정보의 효율적 처리를 위해 소장데이터를 서지데이터에서 분리하여 관리하는 방식을 취하고 있다(노지현, 이은주, 2018). 이외에도, 전국의 1,980개 도서관(2023년 6월 말 기준)에서 수집한 서지레코드를 통합한 KOLIS-NET(국가자료종합목록)이나 OCLC의 WorldCat 종합목록도 '유형 2'의 구조를 기본 구조로 채택하고 있다. 다만, WorldCat에서는 '저작' 단위로 서지레코드를 구축하되, 데이터의 효율적인 공유를 위해 동일 저작이라 하더라도 입력수준(리더/17)이 상이한 서지레코드에 대해서는 중복을 인정하는 정책을 취하고 있다.

서지데이터의 내용적 품질에 있어서도 부산지역 도서관이 구축한 서지데이터에서는 정확성, 완전성, 표준의 준수, 적용방식의 차이 등에서 많은 문제점이 발견되었다. 특히, 자료의 식별과 검색에 결정적인 영향을 주는 자료유형의 구분이나 물리적 형태에 관한 정보, 발행정보 등에서 기본적인 오류가 빈번하게 발견되었고, 이외에도 자료의 특성을 설명하는 정보나 자료의 주제적 속성 등에서 불충분한 데이터나 데이터 기술 방식의 비일관성 등에서 비롯되는 품질 문제도 확인되었다. 이러한 서지데이터의 내용적 품질은 주로 각 도서관이 사용하는 지침의 빈약성과 오류, 표준에 대한 실무진의 무지와 오해, 구축된 데이터에 대한 검수과정의 소홀 등과 같은 원인이 복합적으로 작용한 결과로 확인된다. 따라서 서지데이터의 품질이나 검색의 효용성 측

면에서 결정적인 영향을 미치는 문제점을 최소화하기 위한 실행방안의 마련도 시급한 상황이라 볼 수 있다.

2. 품질 제고를 위한 실행전략의 제안

이번 연구는 '서지데이터의 품질 제고' 뿐만 아니라 '업무효율' 및 '데이터 공유를 위한 기반 구축'을 동시에 실현하는 것을 목표로 하여 실행전략을 도출하였다. 실행전략의 도출을 위해 먼저 3가지 원칙을 설정하였다. 첫째, 새로운 서지구조의 적용을 대비하고, 이와 동시에 서지데이터를 기반으로 하는 부가적 서비스(이용통계, 서평정보, 서지적 관련 자료 안내 등)를 구현하는데 적합한 실행방안을 마련한다. 이를 위한 전제조건은 동일 자원에 대한 서지데이터의 통합 및 동일 자원의 상이한 매체에 대한 명확한 구분이라 할 수 있다. 둘째, 부산지역 도서관들의 의견수렴 결과와 통합DB 구축에 투입되는 노력과 비용에 비해 지역 종합목록의 실질적인 활용도가 낮을 수 있다는 점을 종합적으로 고려하여, 지역 도서관에서 구축한 서지데이터를 인위적으로 통합하고 중복처리 후 고품질의 통합DB를 업무에 활용하는 방안은 배제한다. 대신에, 현 단계에서는 개별 도서관 차원에서 서지데이터의 품질을 개선하기 위한 공동의 방안 도출에 집중하고, 이를 통합DB의 품질 개선을 위한 기반으로 활용한다. 셋째, 실행방안은 각 도서관에서 개별적으로 추진해야 할 실무적 방안이 아니라 지역 도서관이 협력하여 공동으로 대처하기 위한 정책적 방안에 중점을 둔다.

이러한 목표와 원칙을 토대로 이번 연구에서 제안하는 실행전략은 ① 서지데이터의 구축을 위한 거버넌스의 구축, ② 서지데이터의 통합, ③ 지역 기반의 표준 지침의 마련, ④ 직무역량 강화를 위한 실무교육의 강화 등이다. 이들 4가지 전략은 각각 독립된 개별적인 것이 아니라 상호 영향을 미치는 연결된 것이라 할 수 있다. 실행전략이 성공적으로 정착되면 서지데이터의 품질을 전반적으로 개선할 수 있는 것은 물론이고, 외부 데이터의 반입에 대한 절대적인 의존도 또한 개선할 수 있을 것으로 기대된다. 세부내용은 다음과 같다.

지역 도서관에서 서지데이터를 구축하는 가장 '이상적'이고 '효율적'인 방식은 국외(특히 북미 도서관) 사례와 같은 중앙집중식(centralized cataloging) 업무구조라 할 수 있다. 지역 공공도서관 간에 동일 자료의 소장율이 비교적 높은 편이고, 지역 종합목록의 효율적인 구축을 고려할 때에도 각 도서관마다 중복적으로 업무를 수행하기보다는 광역대표도서관 차원에서 서지데이터의 구축을 일괄 수행하는 구조가 가장 비용효과적이기 때문이다. 그러나 중앙관과 분관의 이원화 체제로 통합 운영되는 국외 도서관과 달리, 인력과 예산 등 운영체제가 도서관별로 명확히 구분되어 있는 국내의 상황에서 서지데이터의 구축만을 위한 중앙집중식 업무구조를 만드는 것은 현실성이 매우 떨어진다고 할 수 있다.

이에 대한 대안이 통합DB를 이용한 서지레코드의 공유라 할 수 있다. 즉, 부산지역 도서관의 서지데이터를 한 곳에 통합・저장하여 이를 공동 활용하자는 것인데, 이 방법 또한 앞서 언급하였듯이 도서관의 요구가 높지 않은 상황에서는 비용 대비 효과가 크게 떨어진다고 볼 수 있다. 특히, 외부로부터 서지데이터를 반입하는 도서관이 절대 다수인 상황에서, 이를 일거에 개선할 수 있는 위로부터의 정책적 결단이나 내부에서의 자성적 움직임이 없이는 이 방법 또한 결코 문제해결을 위한 적실한 해법이 될 수 없다.

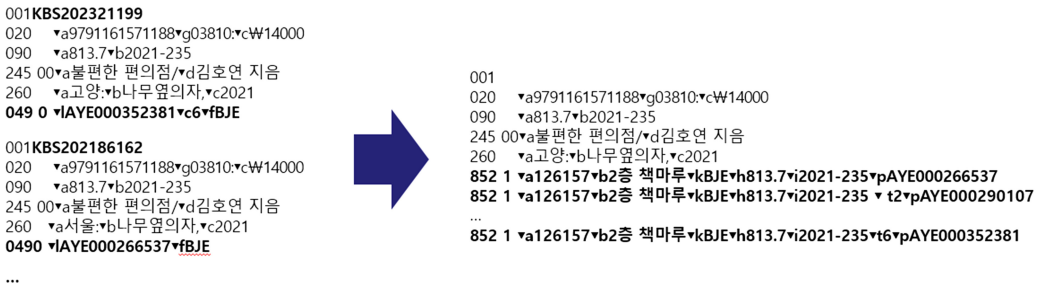
이에 따라 당장 적용가능한 현실적인 방법은 광역대표도서관이 주도가 되어, 실무자 중심의 실질적인 협의체(실무위원회)를 조직하여 개별 도서관의 서지데이터 품질을 지역 차원에서 엄격히 관리해 나가는 것이라 할 수 있다. 앞서 살펴본 품질분석 결과에서, A도서관은 대표도서관이었다. 대표도서관에서 구축한 서지데이터의 품질은 이번 연구에서 다른 도서관에 비해 매우 우수한 편으로 나타났다. 대표도서관은 자체 도서관을 위한 노력뿐만 아니라 이러한 경험을 지역 도서관들과 공유하고 확장해 나가야 하는 역할과 책임을 지니고 있다. 나아가, 부산지역 전체 도서관의 통합검색서비스를 제공하기 위해서도 개별 도서관에서 구축한 서지데이터를 엄격히 제어하는 역할을 책임감 있게 수행할 필요가 있다. 이러한 이유에서, 통합 ILS에 축적된 서지데이터의 품질을 수시로 모니터링하면서 문제점을 공유하고, 해결방안을 지역 차원에서 공동으로 모색하거나 실행하는데 실질적인 역할을 담당할 거버넌스의 구축이 절실히 필요하다고 할 수 있다.

서지데이터의 품질을 '조정' 및 '통제'하기 위한 이러한 거버넌스는 OCLC와 KERIS 등에서도 고품질의 서지데이터를 유지하기 위해 이미 도입하고 있는 제도이기도 하며, 그 효과도 어느 정도 검증된 것이라 할 수 있다. 협의체를 기반으로 업무조정과 데이터의 공유를 위한 기반과 공감대를 충분히 형성하고 동시에 높은 수준의 표준화가 이루어지면, 다음 단계에서 실질적인 통합환경(서지유틸리티의 운영)을 위한 후속논의도 진행할 수 있을 것이다.

두 번째로 공공도서관에서 시급히 개선할 사항은 동일 자원에 대한 서지데이터의 '통합'이라고 할 수 있다. 이는 당장의 업무효율이나 예산절감을 위해서뿐만 아니라 서지데이터를 기반으로 창출되는 다양한 부가적 서비스의 개발이나 새로운 서지구조로의 전환을 위해서도 반드시 개선이 필요하다. 따라서 동일한 자료에 대해서는 하나의 서지데이터를 원칙으로 하되, 복수의 소장정보는 하나의 서지레코드에 통합 기술하거나(〈그림 3〉 참고) 별도의 DB로 관리하는 방식으로 개선되어야 할 것으로 보인다. 이에 서지데이터를 직접 구축하는 도서관에서는 복본에 대해 별도의 서지레코드를 구축하지 않고 기존 레코드에 소장정보만 추가하는 방식으로 업무구조와 시스템의 기능을 개편하고, 외부에서 서지데이터를 반입하여 사용하는 도서관에서는 복본에 대한 서지데이터의 반입을 중단하기 위한 업무절차의 개선을 마련해야 할 것이다.

서지레코드를 통합하고 소장정보를 재구성하는 방안은 2단계로 나누어 추진할 수 있다. 1단계는 신규 레코드를 작성하면서 서지레코드와 소장정보를 분리하는 방식을 곧바로 적용하는 것이고,

2단계는 이미 구축된 소급 데이터를 이와 같은 구조로 재정비하는 방안을 마련하여 적용하는 것이다. 소급 데이터의 적용에 대해서는 아래와 같이 기계적으로 데이터를 분리하거나 추출하는 방법의 적용이 필요할 것으로 보인다.



〈그림 3〉 서지데이터의 통합 및 소장정보 처리 방안

세 번째는 지역 도서관들이 업무에 활용할 「서지데이터 표준 입력지침」을 공동으로 개발하는 방안이다. 현재 부산지역 도서관들은 도서관별로 지침을 작성하여 활용하고 있지만, 주목할만한 내용적 차이가 없고 대부분 유사한 수준으로 작성되어 있다. 특히, 최소수준에 준하는 기본적인 데이터 위주로 되어 있고, 지침에서의 오류나 누락 등의 문제는 모든 도서관이 공통으로 가지고 있다. 따라서 개별 도서관의 지침이 아니라 부산지역 도서관을 위한 '공동' 지침을 개발하여 이를 지역 내 모든 도서관에서 공동으로 활용하도록 하는 것이 업무효율은 물론이고 통합DB의 데이터 품질을 제고하는 데에도 바람직할 것이다. 앞서도 언급했듯이 이는 높은 수준의 데이터 표준화는 물론이고 서지데이터의 품질을 강화하기 위한 수단으로 매우 유용하게 활용될 수 있다.

표준 지침은 광역대표도서관에서 주도하여 운영할 협의체(실무위원회)에서 공동으로 개발하여 관리하되, 정기적인 논의를 통해 지속적으로 추가·보완해 나가는 것이 바람직할 것이다. 표준 입력지침의 개발 과정에서 특별히 고려할 점을 몇 가지 제안하면 다음과 같다.

첫째, 세부지침의 작성에 앞서 신규 데이터의 작성 기준을 명시할 필요가 있다. 앞서 언급하였듯이 동일 자료에 대해 서지 통합을 원칙으로 할 때, 서지데이터가 이미 존재하는 경우에는 소장정보만 추가한다. 다만, 동일 자료인지의 여부가 불확실하고 새로운 서지데이터를 작성해야 하는지 모호할 경우를 대비하여 신규 레코드의 작성 기준을 명확히 설정할 필요가 있다. 앞서 살펴보았듯이 동일 판에 상이한 발행년도를 가진 자료 혹은 동일 자료의 복제본(점자자료 포함)이나 상이한 매체의 경우 등에 대해 신규 서지데이터를 작성하거나 기존 데이터에 통합 기술하기 위한 기준을 명확히 설정한다.

둘째, 현재의 지침에서는 '도서', '비도서', '전자자료'로 구분하여 세부지침을 제시하고 있다. 그

러나 자료유형별로 수록해야 할 지침의 내용이 매우 상이하므로 이를 일반도서, 점자자료, 전자책, 오디오북, 동영상자원 등과 같이 공공도서관에서 실제 소장하는 자료유형 중심으로 세분화하여 지침을 마련할 필요가 있다. 자료유형별 세부지침에는 앞서 살펴본 것처럼 반복적으로 나타나는 문제점을 강조하여 수록하고, 자료유형별로 특별히 고려해야 할 사항이나 부호화정보의 상세한 입력기준을 포함해야 할 것이다.

셋째, KORMARC에 정의된 '필수' 요소 외에 부산지역 도서관에서 공통적으로 적용할 '핵심' 요소를 별도로 설정하여 운용한다. 핵심요소는 필수요소는 아니지만 자료검색이나 식별, 나아가 데이터의 통합이나 새로운 서지구조로의 전환에서 특히 중요하게 사용될 데이터로서 부산지역 도서관들이 공통적으로 반드시 포함해야 하는 데이터를 의미한다. 필요하다면 도서관의 성격(공공도서관, 공립작은도서관)에 따라 핵심요소를 달리 설정하여 운영하는 방안도 고려할 수 있다. 지역 차원에서 핵심요소를 별도로 지정하여 적용하는 경우 KORMARC 리더/17(입력수준)에는 "4"(핵심수준)를 적용한다.

넷째, 동일 자료의 상이한 판이나 자료의 형태적 정보를 명확하게 구분하기 위해 데이터의 일관된 기술방법과 풍부한 예시를 지침에 수록할 필요가 있다. 가령, 자료의 유형이나 물리적 특성을 정확하고 일관된 방식으로 기술하기 위해 리더/06(레코드 유형), 007(형태기술필드), 245 필드의 ▼h(자료유형표시), 300 필드의 ▼a(특정자료종별) 등을 종합하여 적용한 상세한 사례를 지침에 포함한다. 엄격하고 제어된 형태로 작성된 이들 데이터는 향후 336(내용유형), 337(매체유형), 338(수록매체유형) 필드로 확장하거나 새로운 서지구조로 전환할 때 매우 유용하게 활용할 수 있다.

마지막으로, 공공도서관 차원에서 전거데이터의 구축이나 국가전거의 활용이 현실적으로 어려운 상황에서 과도기적으로 접근점(전거형과 이형)을 일관되게 기술하기 위한 방안의 마련도 필요하다. 제어되지 않은 형식이라도 일관성 있게 구축된 접근점은 향후 전거데이터의 자동 생성이나 연결을 훨씬 효율적으로 추진할 수 있을 것이기 때문이다. 따라서 부산지역 도서관들이 적용할 이름전거의 작성 방법을 마련한다.

이처럼 부산지역 도서관을 위한 표준 지침이 만들어지면, 이를 토대로 서지데이터의 품질을 강화하기 위한 실무교육 또는 직무연수를 지역 단위에서 활성화하는 제도적 절차의 마련도 필요하다. 따라서 실행전략의 마지막 제안은 실무교육 또는 직무연수의 강화라 할 수 있다. 표준 지침을 적용하는 과정에서 자주 발견되는 문제점을 중심으로 맞춤식 교육 또는 근거 기반(evidence-based) 교육을 진행한다면 실무교육은 곧 서지데이터의 품질을 강화하기 위한 중요한 수단으로 활용될 수 있다. 이를 위해 교육은 교수진이나 외부 강사보다 표준 지침의 개발에 참여한 협의체(실무위원회) 소속의 실무진이 직접 담당하되, 피교육자의 수준별, 자료유형별 등으로 교육내용을 차별화하여 진행하는 것이 교육의 효과도 높이고, 결과적으로 지역 차원에서 높은 수준의 데이터 표준화를 실행할 수 있는 해법으로 활용될 수 있을 것이다.

V. 결 론

이 연구는 부산지역 도서관에서 구축한 서지데이터의 품질을 분석하고 데이터의 품질제고 및 업무효율을 증대하기 위한 실행전략을 제안하는데 목적을 두고 진행되었다. 연구결과를 요약하면 다음과 같다.

부산지역 49개 공공도서관 중 3개 도서관을 제외한 모든 도서관이 외부로부터 서지데이터를 반입하고 있는 것으로 나타났다. 따라서 부산지역 공공도서관에서 구축된 서지데이터는 대부분 외부에서 만들어진 것이라 할 수 있으며, 반입 데이터도 대부분 동일 기관에서 구축한 것이어서 서지데이터의 내용적 차이가 크게 없다고 볼 수 있다. 그럼에도 불구하고 모든 도서관이 개별적으로 MARC 반입에 비용을 투입하고 있고, 이는 도서관 전체적으로 볼 때 적지 않은 비용이 중복투자 되는 구조라 할 수 있다.

서지데이터의 품질분석 결과에서 가장 근본적인 문제는 동일 자원(복본)에 대해 책 단위로 서지레코드를 중복 생성하는 구조로 나타났다. 이는 데이터의 관리는 물론이고 업무효율을 크게 저하하고 MARC 반입 시 예산낭비의 주요 요인이 되고 있을 뿐만 아니라 향후 전거데이터나 부가적 콘텐츠의 활용, 나아가 새로운 서지모형의 적용에서도 많은 제약으로 작용할 것으로 분석되었다. 서지데이터의 내용적 품질에 있어서도 정확성, 완전성, 표준의 준수, 기술방식의 차이 등에서 많은 문제점이 발견되었고, 이러한 문제가 결과적으로 통합목록으로 그대로 이어지고 있었다. 주요 품질 문제로 자료의 식별과 검색에 결정적인 영향을 주는 자료유형의 구분이나 물리적 형태에 관한 정보, 발행정보 등에서 오류가 빈번하게 발견되었는데, 이러한 서지데이터의 내용적 품질은 도서관에서 직접 구축한 서지데이터보다 외부에서 반입한 데이터에서 상대적으로 높게 나타났다.

한편, 서지레코드의 구조와 내용적 측면에서 참고할 내용을 분석하고자 국내외 주요 도서관의 사례를 분석하고 시사점을 도출하였다. 특히, 복본이나 동일 자원의 다양한 매체를 식별하기 위한 데이터의 기술방식은 물론이고, 일관되고 정확한 데이터의 기술 방법, 이름 전거나 주제명 데이터의 품질 차이 등을 명확히 확인할 수 있었다. 이러한 내용을 기초로 하여, 부산지역 도서관의 서지데이터 품질제고, 업무효율화, 데이터 공유체제 구축을 위한 실행전략을 제안하였다. 실행전략을 제안하기에 앞서, 현행 서지데이터 구축 방식(외부 MARC 반입)에 대한 획기적인 개선 의지 없이는 지역 기반의 서지유틸리티 운영이나 광역대표도서관의 서지데이터를 지역 내 도서관들이 공동으로 활용하는 방안이 현실성이 떨어진다고 판단하여 이번 연구의 제안에서는 이를 배제하였다. 대신에, 개별 도서관 단위에서부터 서지데이터의 구축 방식과 품질을 개선하는 전략을 제안하였다. 구체적인 실행전략은 (1) 부산도서관 중심의 거버넌스 구축, (2) 동일 자원에 대한 서지레코드의 통합, (3) 부산지역 도서관에서 공동 사용할 서지레코드 표준 지침의 개발, (4) 실무교육 또는 직무연수의 강화 등이었다.

이 연구는 지역 도서관 차원에서의 서지데이터 구축 모델을 점검하고, 서지데이터의 품질관리와 함께 업무효율을 확보하기 위한 방안을 도출하는데 목적을 두고 있다. 이번 연구에서 제안한 내용은 (1) 정책적으로 해결해야 하거나, (2) 실무 차원에서 당장 실행해야 하는 것도 있으며, (3) 시스템 기능의 개선이 뒷받침되어야 하는 것도 있다. 또한, 장기적인 관점에서 단계적으로 실행방안을 마련하고 이를 실행하기 위한 공동의 노력을 해나가야 하는 것도 있다. 이번 연구의 결과가 공공도서관의 서지데이터의 품질을 제고하고, 나아가 해결책을 지역 차원에서 공동으로 모색하는데 소중한 정책자료로 활용되기를 기대한다.

참 고 문 헌

국가도서관위원회 (2023). 국가도서관통계시스템. 출처: https://www.libsta.go.kr/main

김선애, 이수상 (2006). KOLIS-NET 종합목록 DB의 품질평가. 한국문헌정보학회지, 40(1), 95-117.

노지현, 이은주 (2018). 국내 도서관의 소장정보 관리 현황과 소장정보용 KORMARC의 개정 방안. 한국도서관・정보학회지, 49(4), 93-117. https://dx.doi.org/10.16981/kliss.49.201812.93

부산도서관 누리집 (2023). 출처: https://library.busan.go.kr/busanlibrary/

박진호 (2018). 도서관의 오픈 데이터 품질측정모델 개발. 정보관리학회지, 35(1), 33-59. https://dx.doi.org/10.3743/KOSIM.2018.35.1.033

양수완 (2020). 도서관 공공데이터의 품질에 관한 연구: 도서관 정보나루의 도서 상세 조회 API를 중심으로. 정보관리학회지, 37(4), 181-206. https://dx.doi.org/10.3743/kosim.2020.37.4.181

이제환 (2002). 공동목록 DB의 품질평가와 품질관리: KERIS의 종합목록 DB를 중심으로. 한국문헌정보학회지, 36(1), 61-89.

한국데이터베이스진흥센터 (2000). 데이터베이스 품질평가 항목. 서울: 한국데이터베이스진흥센터, 정보통신부.

한국정보화진흥원 (2018). 공공데이터 품질관리 매뉴얼 v2.0. 대구: 한국정보화진흥원.

Kyprianos, K., Lolou, E., & Efthymiou, F. (2022). Cataloging quality and the views of catalogers in Hellenic Academic Libraries. Cataloging & Classification Quarterly, 60(8), 683-707.

Snow, K. (2011). A Study of the Perception of Cataloging Quality Among Catalogers in Academic Libraries. Denton: University of North Texas.

• **국한문 참고문헌의 영문 표기**
(English translation / Romanization of references originally written in Korean)

Busan Metropolitan Library (2023). Available: https://library.busan.go.kr/busanlibrary/

Kim, Sun Ae & Lee, Soosang (2006). Quality evaluation of a shared cataloging DB: the case of KOLIS-NET. Journal of the Korean Society for Library and Information Science, 40(1), 95-117.

Korea Database Agency (2000). The Standardization of the Quality Evaluation Item on Database. Seoul: Korea Database Agency.

Lee, Jae Whoan (2002). Quality evaluation and management of a shared cataloging DB: the case of KERIS UNICAT DB. Journal of the Korean Society for Library and Information Science, 36(1), 61-89.

National Information Society Agency (2018). Open Government Data Quality Management Manual v2.0. Daegu: National Information Society Agency.

Park, Jin Ho (2018). Developing an assessment model of library open data quality. Journal of the Korean Society for Information Management, 35(1), 33-59. https://dx.doi.org/10.3743/KOSIM.2018.35.1.033

Presidential Committee on Library and Information Policy of Korea (2023). National Library Statistics System. Available: https://www.libsta.go.kr/main

Rho, Jee-Hyun & Lee, Eun Ju (2018). Current status of holdings data in Korean libraries and proposal for revision of KORMARC format for holdings data. Journal of Korean Library and Information Science Society, 49(4), 93-117. https://dx.doi.org/10.16981/kliss.49.201812.93

Yang, Suwan (2020). Quality diagnosis of library-related open government data: focused on book details API of Data for Library. Journal of the Korean Society for Information Management, 37(4), 181-206. https://dx.doi.org/10.3743/kosim.2020.37.4.181