

상업서지를 이용한 국립중앙도서관 목록데이터베이스 평가

Evaluating the Catalog Database of the National Central Library
Using Trade Bibliographies

장혜란. 상명여자대학교 문헌정보학과

Chang, Hye Rhan, Department of Library & Information Science,
Sangmyung Women's University

Using three existing trade bibliographies, the catalog database of the National Central Library was evaluated. A total of 8,901 books sampled are used as the basis for the measurement. Results indicate about 70% coverage for the in print books, 80% coverage for the books published in 1993, and 40% coverage for the books published in 1994. The coverage analysis according to time lag ranged from 18% to 54%. In the subjects of generalities and linguistics, low coverage was found consistently. Areas for future research are briefly examined.

1. 서론

최근 국립중앙도서관 목록자동화와 목록데이터 제공서비스는 우리나라 도서관의 발전적 측면에서 볼 때 매우 고무적인 일이다. 1994년 6월 현재 국립중앙도서관 목록데이터베이스는 광복이후 수집된 도서 약 67만건을 수록하고 있으며, 당도서관의 열람목록으로 활용됨은 물론 모든 도서관을 대상으로 인쇄카드, 디스켓, 마그네틱테이프, 콤팩트디스크 형태로 혹은 온라인으로 제공되고 있다. 목록자동화를 통하여 국립중앙도서관은 서지를 통정하는 국가대표도서관으로서 위상을 확립하게 되었으며, 데이터를 공급받는 각 도서관들은 자관의 편목과 데이터변환업무를 신속하게 수행할 수 있게 되었고, 국가 전체적으로는 자원의 공동활용과 목록데이터의 표준화를 실현할 수 있게 되었다.

국립중앙도서관은 국내 문헌을 당라적으로 수집하기 위하여 납본제도를 운영하고 있으므로 국내 생산 문헌을 모두 소장하고 있어야 하며, 소장 문헌에 대한 엔트리는 목록데이터베이스에 수록되어야 한다. 한편, 이용자가 원하는 문헌정보를 적시에 제공하고 목록데이터가 타도서관에 유용하기 위하여 데이터베이스는 신간도서를 신속하게 수록하여야 한다. 그러나 국립중앙도서관 목록데이터베이스가 이러한 요구를 어느 정도 만족시키고 있는가에 대한 정확한 정보는 찾아보기 어렵다.

본 연구의 목적은 국립중앙도서관 목록데이터베이스의 포괄성과 최신성을 상업서지를 이용하여 측정함으로써 데이터베이스의 질을 평가하는 데 있다. 이러한 측정치는 또한 국립중앙도서관의 문헌전달력을 측정하거나 향후데이

타베이스 개선을 평가하기 위한 기초자료가 될 것이며, 목록데이터베이스를 이용하는 각 도서관들의 계획수립에 유용한 정보를 산출하게 될 것이다.

2. 연구방법

2.1 상업서지의 선택

본 연구를 수행하기 위하여 다음과 같은 기존의 세 가지 상업서지를 선택, 활용하였다.

유통도서 전반에 관한 포괄성을 측정하기 위하여는 "판매도서목록" 1994년판 주제별 목록을 선택하였다. 이 서지는 한국출판협동조합 가입 출판사들이 1993년말까지 간행, 유통하고 있는 도서를 망라적으로 수록하고 있다.

1993년에 간행된 도서의 포괄성을 측정하기 위하여는 "한국출판연감" 1994년판을 선택하였다. 이 서지는 대한출판문화협회에서 1993년 1년동안 문화체육부에 납본한 도서를 모두 수록하고 있다.

1994년도 신간도서에 대한 포괄성과 최신성을 측정하기 위하여 선택한 서지는 교보문고에서 발행하는 "지구촌책정보"에 실려있는 '새로나온 책' 부분이다. 이것은 월간으로 출판되는데 교보문고에서 지닌달 입수된 국내도서를 수록하고 있다. 본 연구에서는 1994년 발행 도서를 포함하고 있는 3월호 부터 9월호까지를 선택하였으며 서지의 각 엔트리 상단에 있는 도서의 발행일자를 최신성의 측정에 사용하였다.

2.2 표집방법과 절차

"판매도서목록"의 도서전체와 "한국출판연감"중 학습참고서를 제외한 도서전체에 대하여 계통표집방법으로 5%의 표집을 추출하였다. 이 두 가지 서지는 그 망라성, 표집확률, 효율성 면에서 볼 때 표집틀로서 적합하다. "지구촌책정보"에 실린 서지는 신간도서의 모집단으로 판단하기 어려우므로 표본을 추출하지 않고 전체를 조사대상으로 하였다.

위의 서지들은 원칙적으로 도서를 수록하고 있다. 그러나 드물게 표본으로 선정된 자료가 도서가 아닌 시청각자료, 전산매체, 혹은 연속간행물인 경우가 발생하였다. 이 때에는 "한국

출판연감"과 "판매도서목록"에서는 해당 표본과 가장 근접한 도서를 취하였으며, "지구촌책정보"에서는 조사대상에서 제외시켰다. 조사대상의 수는 "판매도서목록"의 2,489 건, "한국출판연감"의 730 건, "지구촌책정보"의 5,682 건, 총 8,901 건이다.

2.3 데이터 수집방법

조사대상 엔트리에 대하여 목록데이터베이스를 탐색하여 포함 여부를 판정하였다. 탐색에는 국립중앙도서관 온라인 목록실의 단말을 이용하였다. 탐색작업은 1994년 11월 10일부터 11월 23일까지 2주간 수행되었으며, 탐색에는 ISBN을 포함하여 가능한 모든 탐색점을 사용하였다. 비소장 자료인 경우에 이차탐색을 수행함으로써 수록여부를 재확인하였다.

3. 데이터분석과 논의

탐색결과 얻은 데이터 서지별 포괄성을 측정하였으며, 서지 각각에 대하여 주제별 분석을 시도하였다. 본 연구에 사용된 서지는 모두 한국십진분류법에 준하여 정리되어 있으며 출판·서점가의 관행에 따라 아동도서는 별도로 구분되어 있다. 주제별 분석에는 기존 서지의 주제구분을 그대로 적용하였다.

3.1 유통도서 전반에 관한 분석

"판매도서목록"으로부터 추출한 표본엔트리에 대한 온라인목록 탐색결과는 (표 1)과 같다. 유통도서전반에 관한 목록데이터베이스의 포괄성은 69.83%로 나타났다. 주제별로는 문학에서 포괄성이 가장 높고 그 다음이 사회과학, 철학, 순수과학의 순서로 되어 있다. 포괄성이 가장 낮은 분야는 총류이며 그 다음이 예술이다.

(표 1) 유통도서전반에 대한 포괄성

주 제	표본수	DB수록건수	포괄성(%)
총 류	68	27	39.71
철 학	120	88	73.33
종 교	75	47	62.67
사회과학	458	359	78.38

순수과학	79	57	72.15
기술과학	336	217	64.58
예 술	188	101	53.72
어 학	178	115	64.61
문 학	606	491	81.02
역 사	135	72	53.33
아동도서	246	164	66.67
계	2,489	1,738	69.83

3.2 1993년도 간행 도서에 대한 분석

“한국출판연감”으로부터 추출한 표본엔트리에 대한 온라인목록 탐색결과는 (표 2)와 같다. 1993년도 간행 도서에 대한 데이터베이스의 포괄성은 79.86%로 나타났다. 유통도서전반보다 약 10% 높은 수치이다. 주제별로는 종교에서 가장 포괄성이 높고 문학, 순수과학, 사

(표 2) 1993년 간행 도서에 대한 포괄성

주 제	표본수	DB수록건수	포괄성(%)
총 류	8	5	62.50
철 학	26	22	78.57
종 교	79	71	89.87
사회과학	120	95	79.17
순수과학	16	14	87.50
기술과학	116	84	72.41
예 술	46	35	76.09
어 학	37	23	62.16
문 학	186	163	87.63
역 사	29	20	68.97
아동도서	65	51	78.46
계	730	583	79.86

회과학의 순서로 되어 있다. 포괄성이 가장 낮은 분야는 어학, 총류, 역사의 순서이다.

3.3 1994년 신간도서에 대한 분석

“지구촌책정보”의 서지에 포함된 엔트리의 온라인목록 탐색결과는 (표 3)과 같다. 신간도서에 대한 포괄성은 40%로 나타났다. 이것은 유통자료전반보다도 훨씬 낮으며 93년도 간행자료의 반정도 수준에 불과하다. 주제별로는 철학, 기술과학, 사회과학, 문학에서 포괄성이 평균보다 약간 높다. 포괄성이 낮은 분야는 아동도서, 어학, 총류의 순서이다.

(표 3) 1994년 신간도서에 대한 포괄성

주 제	표본수	DB수록건수	포괄성(%)
총 류	82	23	28.04
철 학	266	121	45.48
종 교	329	129	39.20
사회과학	1,004	423	42.13
순수과학	107	39	36.44
기술과학	1,297	569	43.89
예 술	411	143	34.79
어 학	223	60	26.90
문 학	1,654	564	40.14
역 사	181	72	39.77
아동도서	128	30	23.43
계	5,682	2,273	40.00

신간도서에 대한 최신성 분석을 수행하였다. 엄격히 말하면, 최신성이란 도서의 출판일자와 데이터베이스 수록일자간의 시간 차이의 측정치이다. 그러나 열람목록을 가지고 탐색한 현 상황에서는 데이터베이스 수록일자를 식별할 수 없다. 따라서 본 연구에서는 최신성을 도서 발행일자와 탐색시점을 가지고 시도하였다. 신간도서 중 발행일자 3월부터 8월까지의 엔트리를 조사대상으로 하였다. 11월 10일부터 탐색

이 수행되었으므로 시차는 약 2개월부터 8개월 까지 허용되었다. 발행일자를 구분하여 분석한 결과가 (표 4)에 나타나 있다.

(표 4) 데이터베이스의 최신성 분석

발행일자	표본수	DB수록건수	포괄성(%)
3월	845	455	53.84
4월	1,069	517	48.36
5월	643	255	39.65
6월	613	177	28.87
7월	488	128	26.22
8월	89	16	17.97

(표 4)를 보면, 포괄성이 시차가 적어짐에 따라 급격히 감소하고 있는 것을 알 수 있다. 발행일자 3월의 도서는 약 54%, 6월의 도서는 약 29%, 8월의 도서는 약 18%만이 데이터베이스에 포함되어 있다.

1993년 발행도서에 대한 비교적 높은 포괄성(약 80%)은 데이터베이스가 질적으로 향상되고 있음을 나타낸다. 이 수치는 또한 문화체육부 납본도서 중 약 20%가 데이터베이스에 누락되어 있음을 나타낸다. 유통자료전반에 대한 포괄성은 1993년 발행도서의 포괄성보다 약 10% 낮다. "판매도서목록"이 1993년도 발행도서 대부분을 포함하고 있다는 점을 간주한다면, 발행년도가 소급될수록 포괄성이 더욱 낮아짐을 알 수 있다. 1994년 신간도서의 낮은 포괄성(40%)과 시차에 따른 감소는 심각한 문제를 제기한다. 전년도(1993년)와 비슷한 수준의 납본율은 예상한다면, 최신성의 문제가 납본지연 혹은 관내 업무처리지연에 기인하는가가 식별되어야 할 것이다. 주제별 분석에서 총류, 어학 분야의 일관성있게 낮은 포괄성은 이 분야의 도서수집에 구조적 문제가 있음을 제시한다.

4. 결론 및 제언

이상과 같은 분석결과 국립중앙도서관 목록 데이터베이스의 포괄성과 최신성이 기대치보다

부족함을 알 수 있다. 이것은 국내문헌의 수집 보존 및 서지통정의 기능이 적절히 수행되고 있지 않으며, 탐색자들이 도서를 발견할 확률이 기본적으로 제한되어 있음을 의미한다. 특히 신간도서에 대한 낮은 포괄성과 최신성은 목록데이터를 공유할 수 있는 집중목록서비스의 가능성에 심각한 문제를 제기한다.

낮은 포괄성에 대한 일차적 요인은 도서관계에서 끊임없이 지적되어 온 바와 같이 불완전한 현행납본제도에 있다. 그러나 국립중앙도서관이 본래의 기능을 다하기 위하여는 도서관 자체에서 보다 능동적인 대처방안이 강구되어야 한다. 납본누락 도서를 발굴하여 납본을 요청하고 현행법의 보상규정을 실행할 수 있는 예산을 확보함으로써 포괄성을 향상시켜야 할 것이다. 이러한 노력은 신간도서는 물론 기간도서에 대하여도 수행되어야 한다.

본 연구의 평가대상은 판매도서에 국한되어 있다. 데이터베이스의 포괄성과 최신성에 대한 완전한 평가를 위하여는 기타 납본대상 자료에 대한 연구가 확대 수행되고, 아울러 누락자료에 대한 분석도 이루어져야 한다. 또한 데이터베이스의 질적 개선을 관찰하기 위하여 계속적으로 평가연구가 수행되어야 한다.

<참고문헌>

- 교보문고. 1994. 지구촌책정보. (12)-(18).
- 국립중앙도서관. 1994. 국립중앙도서관요람.
- 대한출판문화협회. 1994. 한국출판연감.
- 한국출판협동조합. 1994. 판매도서목록.
- Comer, C.H. 1981. List Checking as a Method for Evaluating Library Collections. *Collection Building* 3(3): 26-34.
- Lancaster, F.W. 1991. Indexing and Abstracting in Theory and Practice, Chap 9. Evaluation Aspects. Champaign: Univ. of Illinois Press, 116-137.
- Tenopir, C. 1982. Evaluation of Database Coverage : A Comparison of Two Methodologies. *Online Review* (6): 423-441.