

온라인자료의 기계적 품질계수 측정을 통한 컨스펙터스 운영모형에 관한 연구

A Study on an Operating Model of Conspectus for Online Resources based on Mechanistic Quality Factor

우 진 영(Jin-Yung Woo)*

류 성 열(Sung-Yul Rhew)**

〈 목 차 〉

- | | |
|---|--|
| I. 서 론
1. 연구의 필요성 및 목적
2. 연구의 내용 및 방법 | 2. 온라인자료의 주제별 검색현황 |
| II. 온라인자료의 품질과 관련한 선행연구 | IV. 기계적 품질계수의 측정과 컨스펙터스 운영모형 |
| III. 온라인자료의 이용 및 검색현황
1. 온라인자료의 주제별, 유형별 소장 및 이용현황 | 1. 온라인자료의 기계적 품질계수 측정
2. 온라인자료 컨스펙터스 운영모형 |
| | V. 결 론 |

초 록

본 연구는 기계적 품질계수 개념을 도입하여 온라인자료에 적합한 컨스펙터스 운영모형을 개발하기 위한 것이다. 온라인자료의 주제별 소장율, 이용율, 검색율 등 자료활용행태에 관한 정보를 토대로 주제별 기계적 품질계수를 산출한 후 이를 기존의 컨스펙터스의 주제별 집서수준에 반영하는 방식으로 운영할 것을 제안하였다. 그 결과 이용율이 높은 기술과학과 검색율이 높은 역사와 사회과학 등 주제의 집서수준을 대폭 상향조정하고 여타 주제들은 현행유지 또는 하향조정 되어야 하는 것으로 나타났다.

키워드: 디지털도서관, 온라인자료, 컨스펙터스, 이용계수, 이용자만족, 기계적 품질, 장서개발정책

ABSTRACT

This study is to develop an operating model of Conspectus for online resources by using mechanistic quality factor. Based on some empirical statistics about resource use behaviour such as collection rate, use rate, search rate, the mechanistic quality factor was calculated and applied to each subject by adjusting its acquisition level of Conspectus. according to the result of this study, Acquisition levels of the subject such as technical science which shows higher in use rate, and those such as history and social science which show higher in search rate, should be adjusted to a higher acquisition level, and those of others to the same or a lower level.

Keywords: Digital Library, Online Resource, Conspectus, Use Factor, User Satisfaction, Mechanistic Quality, Collection Development Policy

* 숭실대학교 대학원 IT정책경영학과 박사과정(jinyungwoo@naver.com)(제1저자)

** 숭실대학교 컴퓨터학부 교수(syrhew@ssu.ac.kr)(교신저자)

• 접수일: 2012년 11월 14일 • 최초심사일: 2012년 11월 28일 • 최종심사일: 2012년 12월 26일

I. 서론

1. 연구의 필요성 및 목적

도서관은 현재의 장서수준, 미래의 수집의지와 목표, 보존의지와 강도를 종합한 개념으로서 컨스펙터스(Conspectus) 방법을 사용하고 있다. 일반적으로 컨스펙터스라고 하면 도서관에서 주제별 자료수집 수준을 정해 놓은 것을 의미한다. 현재 우리나라의 디지털도서관은 전통적인 종이책 중심의 장서개발정책에 온라인자료를 포함하여 온라인자료의 수집정책을 수립, 시행하고 있으나, 오프라인자료와 온라인자료의 차이와 특성을 고려한 온라인자료에 맞는 컨스펙터스는 아직 충분히 개발되지 않고 있다.

온라인자료는 오프라인자료와는 자료속성과 이용행태가 크게 다르므로 장서개발정책이나 컨스펙터스 수립, 운영도 달라져야 한다. 디지털도서관이 온라인자료를 확충할 때 모든 주제와 유형의 자료를 동일한 수준으로 수집하는 것보다는 이용율이 높은 자료 등 활용도를 높일 수 있는 자료의 확충에 집중할 필요가 있다. 즉, 디지털도서관 온라인자료 수집정책은 이용자들이 원하는 주제와 유형의 자료를 더 많이 우선적으로 수집, 제공할 필요가 있는 것이다.

본 연구의 목적은 디지털도서관의 온라인자료 수집에 있어서 자료이용 서비스의 기계적 품질을 측정하여 컨스펙터스의 수립과 운영에 반영하기 위한 것이다. 온라인자료의 소장율, 이용율, 검색율 등 기계적 품질과 관련한 변수들을 측정하여 이에 따라 기존의 컨스펙터스상의 집서수준을 수정한다면 자료의 활용도와 이용자의 만족도가 개선될 수 있을 것이다. 즉, 본 연구의 목표는 온라인자료의 기계적 품질계수가 반영된 컨스펙터스 운영모형을 제안하는 것이다.

2. 연구의 내용 및 방법

본 연구의 목적을 달성하기 위하여 디지털도서관 이용통계의 조사분석 방법을 사용한다. 첫째, 문헌연구를 통해 디지털도서관의 온라인자료 수집과 관련한 연구 사례 및 관련통계를 바탕으로 이론적 고찰의 틀을 마련한다. 둘째, 국립중앙도서관 디지털도서관의 소장율, 이용율, 검색율 등의 실제 도서관이용통계 자료 조사분석을 통하여 주제별 기계적 품질을 측정하고 이를 컨스펙터스의 주제별 집서수준에 반영하는 운영모형을 제안한다.

디지털도서관 온라인자료의 기계적 품질 측정 및 컨스펙터스 개선의 가능성을 분석하기 위하여 국립중앙도서관 디지털도서관의 온라인자료 소장 및 이용현황을 조사, 분석하고 이를 기초로 온라인자료의 이용계수, 검색계수 및 기계적 품질계수를 도출하고자 한다. 국립중앙도서관 디지털도서관

관을 선택한 이유는 국립중앙도서관은 국내 최대의 장서를 소장하고 있는 국가대표도서관인 동시에 우리나라의 디지털자료운영정책을 총괄, 선도하고 있는 국내 최대 디지털도서관이기 때문이다.

국립중앙도서관의 컨스펙터스는 1(최소수준), 2(기본정보수준), 3(학습교육지원수준), 4(연구수준), 5(망라적 수준) 등 5단계의 집서수준을 구분 적용하고 있는데, 온라인자료는 3(학습교육지원수준)으로 설정하고 있다. 이는 국립중앙도서관이 설정한 컨스펙터스 내에서 가장 낮은 단계이다. 현재 시행되고 있는 국립중앙도서관의 컨스펙터스는 <표 1>과 같다.

<표 1> 국립중앙도서관의 컨스펙터스

주제	아날로그자료		디지털자료			
	국내자료	외국자료	국내자료		외국자료	
			오프라인	온라인	오프라인	온라인
총류	5	3	5	3	3	3
철학	5	4	5	3	4	3
종교	5	3	5	3	3	3
사회과학	5	4	5	3	4	3
자연과학	5	4	5	3	4	3
기술과학	5	4	5	3	4	3
예술	5	4	5	3	4	3
언어	5	4	5	3	4	3
문학	5	4	5	3	4	3
역사	5	4	5	3	4	3

* 국립중앙도서관 홈페이지(www.nl.go.kr) 도서관소개/ 도서관업무/ 장서개발정책/ 자료유형별 장서개발지침 참조.

II. 온라인자료의 품질과 관련한 선행연구

디지털도서관의 온라인자료 이용실태와 장서의 평가를 연계시킨 국내의 선행연구는 아직 종이책 중심의 전통적 도서관 장서개발정책의 한 부분으로서 전자자원의 확충과 관리 전반에 대한 개론적인 연구와, 디지털도서관의 서비스 개선을 위한 이용자 요구분석 등의 실천적인 연구가 일부 나오고 있으나, 아직 온라인자료의 유형별, 주제별 이용실태나 자료수집정책과 관련한 심도있는 연구는 많지 않은 실정이다.

이진경과 이지연은 공공도서관 이용자들을 대상으로 도서관자료의 유형별 주제별 분류표를 제시하고 유형별, 주제별 이용희망자료를 조사하여 이용자들이 가장 이용하고 싶은 유형별 자료는

4 한국도서관·정보학회지(제43권 제4호)

소설이며 주제별 자료는 문학이라는 조사결과를 발표했다.¹⁾ 두 사람의 연구는 도서관자료에 대한 주제별 선호도의 차이를 파악하려는 유익한 시도이기는 하나 온라인자료가 아닌 오프라인자료 연구라는 점에서 온라인자료에 적용하기에는 한계가 있다.

장덕현은 장서개발정책에 있어서의 컨스펙터스 적용의 중요성을 주장하면서 종이책을 위주로 한 기존의 컨스펙터스 개념을 전자자원까지 확대하는 수정된 개념의 컨스펙터스 모형을 제안하였다. 장덕현은 기존의 컨스펙터스 집서수준분류를 전자자원에 적합하게 세분류하는 데에 중점을 두었다는 점에서 기존 분류의 틀 안에서 주제별 평가방안을 제안한 본 연구와 대비된다.²⁾

박진희는 컨스펙터스방법을 이용하여 전문도서관의 장서를 평가하기 위하여 주로 소장율, 주문율 등의 양적 데이터에 기초한 목표장서수준을 산출하였으나 이 역시 오프라인자료에 대한 연구로서 이용율, 검색율 등의 데이터는 사용되지 않았다.³⁾

윤희윤과 김일영은 Bonn의 이용계수(use factor)개념을 적용하여 장서구성비율과 이용실적을 연계한 장서평가방법을 제안하였다.⁴⁾ 특히 윤희윤과 김일영의 연구는 비록 오프라인자료를 대상으로 하였으나 장서평가에 있어서 주어진 소장자료에 대한 이용자의 주관적 평가에 의존하던 기존의 평가방법에서 벗어나 소장과 이용 등 실증적 분석에 대한 관심을 불러 일으켰다는 점에서 도서관자료 평가 연구에 중요한 시사점을 던져 주었다.

이정수와 남영준은 행위이론을 적용하여 이용자의 맥락을 분석하였는데, 도서관 이용자의 대표적인 행위로 학습 및 정보수집의 두 가지 유형을 제시하였다.⁵⁾ 본 연구에서는 이 두 유형을 중심으로 이용계수와 검색계수의 개념을 도입하여 기계적 품질계수 개념을 구성하였다.

국립중앙도서관의 온라인자료 확충과 관련한 연구로는 국립중앙도서관의 디지털도서관개관 이후의 NDL 정보광장 중장기 발전 방안 연구,⁶⁾ 디지털도서관 서비스만족도 증진을 위한 평가지표 개발연구,⁷⁾ 국립중앙도서관 소장자료 디지털화 중장기 전략 연구⁸⁾ 등이 있으나, 대부분 온라인자료의 확충방향과 서비스개선 관련 개괄적인 분석으로 주제별 자료품질까지 세분화된 연구는 이루어지지 못하였다.

기존의 연구들은 주로 오프라인 도서관자료를 중심으로 도서관평가 또는 장서평가 측면에서의

1) 이진경, 이지연, “이용자 장서이용행태 분석을 통한 공공도서관의 장서개발에 관한 연구,” 한국비블리아학회지, 제22권, 제3호(2011. 9), pp.162-163.

2) 장덕현, “전자자원선정을 위한 컨스펙터스 수정모형연구,” 한국비블리아학회지, 제20권, 제2호(2009. 6), pp.41-43.

3) 박진희, “컨스펙터스방법을 이용한 장서평가연구,” 정보관리학회지, 제15권, 제2호(1998. 12), pp.172-174.

4) 윤희윤, 김일영, “이용계수를 적용한 공공도서관 소설장서 평가,” 정보관리연구, 제42권, 제4호(2011. 12), pp.175-177.

5) 이정수, 남영준, “행위이론을 적용한 도서관 이용자 컨텍스트 정보의 CABAT 모델링,” 한국도서관·정보학회지, 제43권, 제1호(2012. 3), pp.15-17.

6) 남영준 등, NDL 정보광장 중장기 발전 방안 연구(서울 : 국립중앙도서관, 2009), pp.174-176.

7) 정현태 등, 디지털도서관 서비스만족도 증진을 위한 평가지표 개발연구(서울 : 도서관연구소, 2011), pp.80-83.

8) 남영준 등, 국립중앙도서관 소장자료 디지털화 중장기 전략 연구(서울 : 국립중앙도서관, 2011), pp.93-97.

서비스품질이나 이용자 만족도와 관련된 다양한 항목을 종합적으로 측정하였기 때문에 실제로 디지털도서관에서 가장 중요한 요소인 주제별 온라인자료 이용실태에 대한 심층적인 분석과 논의가 이루어지기 어려웠다. 다행스럽게도 최근 장서의 소장과 이용을 연계분석한 몇 개의 연구들과 전자자료의 컨스펙티스에 관한 연구들이 나오고 있는 것은 온라인자료 관련 연구의 발전을 위한 좋은 전기가 되고 있다.

본 연구는 도서관, 디지털도서관, 온라인서비스 등 기관으로서의 도서관에 대한 종합평가가 아니라 디지털도서관의 핵심인 온라인자료의 품질 평가에 초점을 맞춘 것이며, 그 중에서도 주관적, 인간적 품질 보다는 객관적, 기계적 품질을 중심으로 자료의 활용도와 이용자 만족도를 높이는 방안을 제시하고자 한다.

Ⅲ. 온라인자료의 이용 및 검색현황

1. 온라인자료의 주제별, 유형별 소장 및 이용현황

먼저 국립중앙도서관의 온라인자료 소장 및 이용현황을 주제와 유형 측면에서 살펴보면, 2011년 말 기준으로 국립중앙도서관이 소장하고 있는 온라인자료는 총 5,901,901건이며 이용회수는 288,882회로서 4.8%의 소장이용비율을 보이고 있다. <표 2>는 2011년 말 기준 국립중앙도서관의 온라인자료 주제별 소장 및 이용현황이다. 이용현황은 통계자료수집 편의상 2011. 4~2012. 3의 1년간 이용현황을 분석대상으로 하였다. 자료의 소장율이 높은 주제는 기술과학(25.8), 사회과학(11.4), 예술(3.5) 순으로, 이용율이 높은 주제는 기술과학(37.7), 사회과학(10.2), 예술(3.0) 순으로, 소장이용비율이 높은 주제는 기술과학(7.1), 문학(4.8), 총류(4.8), 역사(4.7), 철학(4.6), 사회과학(4.3) 순으로 나타났다.

이용율은 온라인자료의 주제에 대한 이용자의 상대적 관심도를 나타내는 지표이다. 이용자들이 가장 많이 이용한 주제는 기술과학으로 소장율(25.8%)에 비하여 이용율(37.7%)이 높은 편이다. 다음으로 많이 이용한 주제는 사회과학으로 소장율(11.4%)과 이용율(10.2%)이 비슷한 수준이다. 그 밖의 주제들은 이용율이나 소장율 모두 1~3%대에 머물고 있다. 소장이용비율이 가장 높은 기술과학은 소장콘텐츠의 활용도가 매우 높음을 알 수 있으며, 이에 비하여 종교, 자연과학, 언어 등의 주제는 소장콘텐츠의 활용도가 낮음을 알 수 있다.

이용계수는 Bonn이 제안한 지표인데 장서구성과 이용실적을 연계한 자료 활용도를 보여준다. 이용계수가 1보다 크면 소장율에 비하여 이용율이 높은 것으로서 자료의 활용도가 높다는 것을 의미한다. 국립중앙도서관 온라인자료의 주제별 이용계수를 살펴보면, <표 2>와 같이 기술과학은 이

6 한국도서관·정보학회지(제43권 제4호)

용계수가 1.46으로서 가장 높게 나타났고, 다음으로 문학이 1.00으로 나타났다. 이용계수가 1에 가까운 역사(0.95), 총류(0.95), 철학(0.93)도 소장율에 상응하는 이용율을 보이고 있다.

〈표 2〉 국립중앙도서관 온라인자료의 주제별 소장 및 이용현황

구분	소장[A]	소장율(%)	이용[B]	이용율(%)	소장이용비율(%) [B/A]	이용계수
총류	130,529	2.2	6,312	2.1	4.8	0.95
철학	92,046	1.5	4,312	1.4	4.6	0.93
종교	118,831	2.0	3,704	1.2	3.1	0.60
사회과학	673,885	11.4	29,476	10.2	4.3	0.89
자연과학	169,789	2.8	4,454	1.5	2.6	0.53
기술과학	1,526,036	25.8	109,103	37.7	7.1	1.46
예술	212,001	3.5	8,681	3.0	4.0	0.85
언어	89,595	1.5	2,929	1.0	3.2	0.66
문학	145,443	2.4	7,122	2.4	4.8	1.00
역사	144,640	2.4	6,930	2.3	4.7	0.95
기타	2,599,106	44.0	105,859	36.6	4.0	0.83
계	5,901,901	100	288,882	100	4.8	

※ 이 자료는 국립중앙도서관의 온라인자료 소장/이용통계를 근거로 작성한 것임
 (소장통계기준시점은 2011.12.31이며 이용통계는 11.4~12.3을 대상기간으로 하였음).
 ※ 이용계수는 (주제별 이용수/ 전체 이용수)/(주제별 자료수/ 전체 자료수)로 계산하였음.

〈표 3〉은 2011년말 기준 국립중앙도서관의 온라인자료 유형별 소장 및 이용현황이다. 이용현황은 통계자료수집 편의상 2011.4~2012.3의 1년간 이용현황을 분석대상으로 하였다. 소장율이 높은 유형은 전자저널(44.9), 전자책(35.5) 순으로, 이용율이 높은 유형은 전자저널(69.6), 전자책(20.7) 순으로, 소장이용비율이 높은 유형은 전자저널(8.0), 동영상자료(7.2), 사운드(5.0), 전자책(3.0) 순으로 나타났다. 이용계수가 높은 유형은 장애인대체자료가 6.50으로 현저하게 높았고 전자저널(1.55), 동영상자료(1.00), 사운드(1.00) 등으로 나타났다.

이용율이 가장 높은 유형은 전자저널로서 소장율(44.9%)에 비하여 이용율(69.6%)이 높은 편이다. 그 다음으로 많이 이용한 유형은 전자책이나 소장율(35.5%)에 비하면 이용율(20.7%)은 낮은 편이다. 그 밖의 유형은 소장율이나 이용율 모두 낮은 수준이다. 소장이용비율이 높은 장애인대체자료, 전자저널, 동영상자료, 사운드 등이 소장자료의 활용도가 높게 나타났고, 이에 비해 전자책, 이미지, 웹사이트 등은 상대적으로 소장자료의 활용도가 낮게 나타났다. 기록물, 소프트웨어, 데이터세트, 콜렉션 등의 유형은 극히 낮거나 없는 실정이다. 동영상이나 사운드는 소장율은 아직 미흡하지만 소장이용비율은 비교적 높은 편으로 콘텐츠확보시 활용도가 높음을 알 수 있다.

〈표 3〉 국립중앙도서관 온라인자료의 유형별 소장 및 이용현황

구분	소장[A]	소장율(%)	이용[B]	이용율(%)	소장이용비율 (%) [B/A]	이용계수
전자책	2,100,182	35.5	63,824	20.7	3.0	0.58
전자저널	2,655,163	44.9	214,090	69.6	8.0	1.55
동영상자료	1,236	0.02	89	0.02	7.2	1.00
이미지	124,355	2.1	2,966	0.9	2.3	0.42
사운드	1,908	0.03	96	0.03	5.0	1.00
소프트웨어	0	0	0	0	-	-
데이터세트	0	0	0	0	-	-
컬렉션	0	0	0	0	-	-
웹사이트	132,397	2.2	1,440	0.4	1.0	0.18
장애인대체자료	14,474	0.2	4,026	1.3	27.8	6.50
기록물	28,305	0.4	14	0.004	0.04	0.01
기타	843,881	14.2	20,854	6.7	2.4	0.47
계	5,901,901	100	307,413	100	5.2	

※ 이 자료는 국립중앙도서관의 온라인자료 소장/이용통계를 근거로 소장율, 이용율, 소장이용비율을 계산하여 작성하였음.

유형별 이용계수를 살펴보면, 장애인대체자료의 이용계수가 6.50으로 가장 높게 나타났고, 다음으로 전자저널(1.55), 동영상자료(1.00), 사운드(1.00) 순으로 나타났다. 이용계수가 낮은 유형은 기록물(0.01), 웹사이트(0.18) 등이다. 소프트웨어, 데이터세트, 컬렉션 등의 유형은 소장 및 이용이 없는 유형이다.

그런데 국립중앙도서관의 온라인자료 중에는 오프라인 자료를 디지털화한 자료를 의미하는 디지털원문이 포함되어 있는데 이 디지털원문은 대부분 전자책과 전자저널의 두 유형에 속하며 일부 자료는 이미지(지도)와 기록물에 속한다. 디지털원문은 국립중앙도서관이 소장하고 있는 방대한 오프라인자료를 디지털자료로 전환한 것으로서 이용자들이 국립중앙도서관에 오지 않고도 귀중한 장서들을 이용할 수 있게 해준다는 면에서 디지털도서관의 장점과 특징을 보여준다. 국립중앙도서관 디지털원문 온라인자료의 소장 및 이용현황을 살펴보면 〈표 4〉와 같다.

소장율이 높은 콘텐츠는 관보(25.6), 단행자료(23.5), 신문(18.0), 일본어자료(11.1), 귀중본(7.5), 고서(5.0) 순으로, 이용율이 높은 콘텐츠는 고서(38.4), 단행자료(20.5), 한국관련외국어자료(9.3), 관보(7.4) 순으로 나타났다. 그리고 소장이용비율이 높은 콘텐츠는 한국고전백선(3456.6), 한글판고전소설(1371.7), 우수학술도서(1067.5), 한국관련외국어자료(611.7), 고서(525.5), 문화관광부발간자료(397.1), 한국고전적자료(317.1), 독도관련자료(301.7) 순으로 나타났다. 그리고 이용계수가 높은 콘텐츠는 한국고전백선(50.00), 한글판고전소설(24.00), 우수학술도서(14.00), 한국관련외국어자료(9.30), 고서(7.68), 독도(6.00), 문화관광부발간자료(5.50), 한국고전적자료(5.33) 등으로 디지털원문의 이용계수가 매우 높은 것으로 나타났다.

8 한국도서관·정보학회지(제43권 제4호)

〈표 4〉 국립중앙도서관 온라인자료 중 디지털원문 소장 및 이용현황

구분	소장[A] (2011. 12기준)	소장율 (%)	이용[B] (2011. 1~12)	이용율 (%)	소장이용비율 (%) [B/A]	이용계수
한국관련외국어자료(1945년 이전)	7,725	1.0	47,261	9.3	611.7	9.30
한국고전백선	83	0.01	2,869	0.5	3456.6	50.00
대학기본학습서	2,577	0.3	1,123	0.2	43.5	0.66
한국고전적자료	2,628	0.3	8,334	1.6	317.1	5.33
인문과학박사학위논문	7,534	1.0	2,690	0.5	35.7	0.50
귀중본-기사	54,903	7.5	18,044	3.5	32.8	0.46
신문-기사(면수)	132,137	18.0	24,604	4.8	18.6	0.26
문화관광부발간자료	1,524	0.2	6,053	1.1	397.1	5.50
한글판 고전소설	905	0.1	12,414	2.4	1371.7	24.00
독도관련자료	115	0.01	347	0.06	301.7	6.00
고서	36,955	5.0	194,029	38.4	525.0	7.68
우수학술도서	37	0.005	395	0.07	1067.5	14.00
신문학100년 대표소설	497	0.06	1,236	0.2	248.6	3.33
해외수집기록물	20,875	2.8	17,333	3.4	83.0	1.21
고지도	90	0.01	66	0.01	73.3	1.00
일본어자료(1945년 이전)	81,853	11.1	16,298	3.2	19.9	0.28
관보-기사	187,710	25.6	37,866	7.4	20.1	0.28
정부간행물	12,166	1.6	4,794	0.9	39.4	0.56
단행자료	172,457	23.5	103,724	20.5	60.1	0.87
국내발간한국관련외국어자료	234	0.03	45	0.008	19.2	0.26
어린이청소년관련자료	8,450	1.1	5,422	1.0	64.1	0.90
합계	731,455	100	504,947	100	69.0	

※ 이 자료는 국립중앙도서관의 디지털원문 온라인자료 소장/이용통계를 근거로 소장율, 이용율, 소장이용비율을 계산하여 작성하였음.

중요한 것은 디지털원문DB전체의 소장이용비율(69%)이 온라인자료전체의 소장이용비율(5%대)보다 10배 이상 높게 나타나고 있는 점이다. 특히 비록 소장율이 낮은 경우이기도 하지만 한국고전백선(3456.6%/ 50.00), 한글판고전소설(1371.7%/24.00), 우수학술도서(1067.5%/14.00) 등은 1000~3000%의 높은 소장이용비율과 14.00~50.00의 높은 이용계수를 나타냄으로써 활용도가 매우 높음을 알 수 있다.

단, 디지털원문의 소장 및 이용 통계와 온라인자료 전체의 소장 및 이용 통계는 DB구축 및 등록 집계방식의 차이 등으로 인하여 소장 및 이용 건수 등 모든 통계값을 일대일 등간척도로 비교하기에는 한계가 있어, 이용율, 이용계수 등 자료이용행태만 비교하였다.

2. 온라인자료의 주제별 검색현황

국립중앙도서관의 디지털도서관 포털인 디브리리상 검색창에 이용자들이 입력한 검색어는 이용자들이 원하는 자료의 내용을 시사해 주는 중요한 정보이다. 국립중앙도서관에서는 매년 이용자들의 검색어 중 상위 1,000개의 검색어와 검색회수 통계를 관리하고 있다. 본 연구에서는 이 중 상위 100대 검색어를 표본으로 하여 주제별 검색율을 추정하고자 한다. 2011년도 국립중앙도서관의 주요 검색어, 검색회수 및 관련 자료현황은 <표 5>와 같다.

<표 5> 2011 상위100대 검색어 및 관련 자료현황

순위	검색어	COUNT	A온라인자료	B디지털원문	B/A(%)
1	한국	52,318	804,886	454,412	56%
2	교육방법	13,098	1,314	338	26%
3	토지	9,324	10,484	8,110	77%
4	무예도보통지	6,651	19	17	89%
5	논문	4,303	83,539	208,920	250%
6	상황인지 추천	4,232	9,619	51	1%
7	북한	4,051	24,903	21,705	87%
8	principles of economic	3,313	1	0	0%
9	1	2,917	226,329	35,795	16%
10	삼국지	2,917	248	161	65%
11	만인보	2,914	10	11	110%
12	도서관	2,814	5,114	4,683	92%
13	다문화	2,693	1,062	365	34%
14	동의보감	2,526	111	78	70%
15~94	(중략)	(중략)	(중략)	(중략)	(중략)
95	홍길동전	648	206	194	94%
96	박경철	644	93	61	66%
97	마음	639	2,744	2,334	85%
98	영성	636	1,985	1,212	61%
99	김진명	635	50	30	60%
100	여행	634	8,170	2,621	32%
	계	203,514	3,334,695	2,583,498	77%

2011년 국립중앙도서관 이용자들이 검색을 위해 입력한 검색어 중 상위 100대 검색어와 관련된 온라인자료 및 디지털원문을 보면, 위 100대 검색어의 검색회수를 모두 합하면 총 203,514회이며, 검색어 1위인 '한국'은 52,318회이고 100위인 '여행'은 634회이다. 그리고 이용자들이 검색한 자료

의 77%가 디지털원문 형태로 소장되어 있음을 보여준다(이 비율은 등록집계 방식상의 차이가 있을 수 있음). 이는 온라인자료 중 특히 디지털원문이 이용자들의 검색요구를 충족하는 내용을 많이 담고 있다고 해석될 수도 있고, 또한 이용자들이 종이책에 있는 내용을 디지털방식에 의해 이용하려는 요구가 크다고도 해석될 수 있다.

위의 검색어와 검색회수를 토대로 온라인자료의 주제별 검색율을 산출하기 위하여 객관적인 방법과 주관적인 방법을 병행하고자 한다. 먼저 객관적인 분석을 위하여 각 검색어 관련 자료의 주제별 분포내역 중 상위3개 주제의 자료수의 비율로 환산, 배분(상위3개 주제까지 환산한 검색회수 합계는 91,930회이며 총 검색회수와 차이는 하위주제 및 기타 분야임)하는 방법으로 주제별 검색회수를 추정할 수 있다. 검색회수가 많은 주제는 사회과학이 28,613회, 기술과학이 24,498회 순으로 나타났고, 검색회수가 적은 주제는 순수과학 232회, 종교 614회 순으로 나타났다. 2011 상위100대 검색어의 검색회수를 주제별로 환산, 배분한 내역은 <표 6>과 같다.

<표 6> 2011 상위 100대 검색어의 검색회수를 주제별 자료수로 환산한 객관적 분석

주제	총류	철학	종교	사회과학	자연과학	기술과학	예술	언어	문학	역사
검색회수	2,916	2,956	614	28,613	232	24,498	8,719	3,260	10,443	9,679
비율(%)	3.1	3.2	0.6	31.1	0.2	26.6	9.4	3.5	11.3	10.5

그런데 위 100대 검색어를 상위3대 주제별 자료수 비율로 환산 배분한 객관적 추정자료의 한계를 보완하기 위하여 검색어에 대한 주관적 분석을 병행하였다. 분석방법은 위 검색어를 먼저 주제와 유형의 두 그룹으로 대분류하고 유형그룹을 걸러낸 후 주제그룹의 검색어들을 다시 이용자의 검색취지에 비추어 Shatford가 분석한 세 가지 정보요구유형인 특정, 보편, 추상의 3개 그룹으로 소분류하여 이에 가장 적합하다고 판단되는 주제를 주관적으로 부여한 결과는 <표 7>과 같다.

<표 7> 2011 상위 100대 검색어의 검색회수를 가장 가까운 주제로 해석한 주관적 분석

주제	총류	철학	종교	사회과학	자연과학	기술과학	예술	언어	문학	역사
검색회수	13,747	13,985	0	32,946	824	9,260	12,451	2,564	27,962	76,001
비율(%)	7.2	7.3	0	17.3	0.4	4.8	6.5	1.3	14.7	40.0

100대 검색어에 대한 주관적 주제부여에 따른 주제별 검색회수 분포를 보면, 검색회수가 많은 주제는 역사(76,001회), 사회과학(32,946회), 문학(27,962회) 순이며, 검색회수가 적은 주제는 종교(0회), 자연과학(824회), 언어(2,564회) 순으로 나타났다. 객관적 분석은 주제별 검색회수가 검색대상 자료의 소장율에 비례하는 특성을 보인 반면, 주관적 분석에서는 검색의 내용을 더 뚜렷하게 알 수 있었다. 검색내용은 주로 한국의 역사, 문화, 경제발전, 한류 등 한국에 대한 관심을 반영하고 있다고 추정된다. 국립중앙도서관 소장 온라인자료의 주제별 검색빈도분포를 보다 더 합리적으로 분석하기 위해 위 두 가지 검색회수 추정치를 합산하여 주제별 검색율을 계산하면 <표 8>과 같다.

<표 8> 2011 상위 100대 검색어의 주제별 검색율 추정

주제	총류	철학	종교	사회과학	자연과학	기술과학	예술	언어	문학	역사
객관적 검색율	3.1	3.2	0.6	31.1	0.2	26.6	9.4	3.5	11.3	10.5
주관적 검색율	7.2	7.3	0	17.3	0.4	4.8	6.5	1.3	14.7	40.0
주제별 검색율	5.1	5.2	0.3	24.2	0.3	15.7	7.9	2.4	13.0	25.2

위의 주제별 검색율과 검색율평균을 대비하여 주제별 검색계수를 계산하면 <표 9>와 같다. 검색계수가 높은 주제는 <표 9>와 같이 역사(2.54), 사회과학(2.44), 기술과학(1.59), 문학(1.31) 순으로, 검색계수가 낮은 주제는 자연과학(0.03), 종교(0.03), 언어(0.24) 순으로 나타났다.

<표 9> 주제별 검색율과 평균값을 비교하여 도출한 주제별 검색계수

주제	총류	철학	종교	사회과학	자연과학	기술과학	예술	언어	문학	역사
검색율(A)	5.1	5.2	0.3	24.2	0.3	15.7	7.9	2.4	13.0	25.2
검색율평균(B)	9.9	9.9	9.9	9.9	9.9	9.9	9.9	9.9	9.9	9.9
검색계수(A/B)	0.51	0.52	0.03	2.44	0.03	1.59	0.79	0.24	1.31	2.54

IV. 기계적 품질계수의 측정과 컨스펙터스 운영모형

1. 온라인자료의 기계적 품질계수 측정

지금까지 온라인자료의 주제별 이용계수와 검색계수 등으로 대표되는 자료이용행태와 관련한 주요 변수들을 살펴보았다. 온라인자료의 품질과 관련한 일반적인 도서관 평가지표를 살펴보면 대체로 도서관경영, 인적자원, 정보자원, 예산 및 시설자원, 정보이용, 정보서비스 등을 들 수 있다. 본 연구는 도서관에 대한 기관 평가가 아니라 정보자원에 대한 평가가 목적이며, 정보자원 평가에 있어서도 얼마나 좋은 자료를, 얼마나 많이 확보하고 있는가에 대한 소장(stock)평가 개념이 아니라, 어떤 책을 수집할 것이냐에 대한 수집(flow개념)평가 개념을 채택하고 있다. 정보자원의 평가 지표 중에서도 소장장서의 수량 등의 지표는 도서관에 대한 평가에는 필요할지 몰라도 수집의 선택에는 큰 관련성이 없다. 또한 정보이용에 대한 평가지표에 있어서도 도서관 전체의 이용율보다는 주제별 이용율 등의 지표가 수집의 선택을 위해 필요한 정보가 된다.

따라서 온라인자료의 수집을 위한 선택의 기준을 제시하는 평가지표로서의 기계적 품질은 앞에서 살펴본 행위이론을 통한 이용자 맥락분석에서 집약된 정보수집 및 학습행위가 가장 중요한 요소라고 볼 수 있으며, 이 요소들이 자료의 주제별로 어떠한 차이를 나타내고 있는가를 조사하여 주제별 집서수준 결정에 반영할 필요가 있는 것이다. 본 연구에서는 학습행위와 정보수집행위를 이용과 검색이라는 두 가지 요소로 변환하여 적용하였다. 왜냐하면 정보수집행위 역시 학습과 검색이라는 두 가지 행위의 결합으로 구성되어 있는 것이므로 결국은 이용과 검색이 두 가지 핵심적인 요소라고 볼 수 있기 때문이다.

이에 따라 온라인자료의 기계적 품질을 계량화하기 위하여 이용계수와 검색계수를 합산, 평균한 기계적 품질계수를 도출하였다. 이 품질계수는 온라인자료의 소장율, 이용율, 검색율이 포함된 주제별 장서평가지표로 볼 수 있다. 이렇게 분석한 <표 10>을 살펴보면, 기계적 품질계수가 높은 주제는 역사(1.75), 사회과학(1.66), 기술과학(1.53) 순으로 나타났고, 기계적 품질계수가 낮은 주제는 자연과학(0.28), 종교(0.31), 언어(0.45) 순으로 나타났다.

<표 10> 주제별 이용계수와 검색계수를 이용한 온라인자료의 주제별 기계적 품질계수

주제	총류	철학	종교	사회과학	자연과학	기술과학	예술	언어	문학	역사
이용계수(B)	0.95	0.93	0.60	0.89	0.53	1.46	0.85	0.66	1.00	0.95
검색계수(C)	0.51	0.52	0.03	2.44	0.03	1.59	0.79	0.24	1.31	2.54
기계적품질계수(D)	0.73	0.72	0.31	1.66	0.28	1.53	0.82	0.45	1.15	1.75

2. 온라인자료 컨스펙터스 운영모형

컨스펙터스는 주제별 집서수준에 대한 도서관의 의지를 표명하는 것이므로 현행 컨스펙터스에 주제별 기계적 품질계수를 반영하여 온라인자료 컨스펙터스를 운영하면 온라인자료의 활용도와 이용자의 만족도가 높아질 것이다

국립중앙도서관의 장서개발정책에 의하면 온라인자료의 집서수준(컨스펙터스Conspectus)은 10개의 주제 모두 3단계 학습교육지원수준(Study or Instructional Support Level)으로 획일화 되어 있다. 이러한 수준설정은 국가의 지식문화유산의 보존, 이용을 책임지고 있는 국립중앙도서관의 임무에 비추어 최소한의 수준을 영구적으로 보장하기 위한 취지라고 이해할 수 있다. 그러나 위에서 본 바와 같이 주제별 이용율이 높은 기술과학과 문학 그리고 주제별 검색율이 높은 역사와 사회과학 등 이용자들이 특별히 더 선호하는 주제는 그에 상응하는 가중치를 부가하여 4단계 연구수준(Research Level) 또는 5단계 망라적 수준(Comprehensive Level)까지 허용할 필요가 있다.

〈표 11〉은 2011년의 주제별 기계적 품질계수를 현행집서수준에 반영한 2012년 온라인자료 컨스펙터스 운영안을 제안한 것이다. 현행집서수준과 기계적 품질계수를 곱하여 수정안을 제시하였기 때문에 집서수준을 소수점 이하 두 자리까지 표기하였다. 기계적 품질계수가 1보다 큰 역사, 사회과학, 기술과학, 문학 등은 집서수준을 현행보다 상향 조정하였고, 기계적 품질계수가 1보다 작은 예술, 총류, 철학, 언어, 종교, 자연과학 등은 집서수준을 현행보다 하향조정하였다. 최종집서수준은 편의상 소수점이하를 반올림하여 현행과 대비하였다.

〈표 11〉 주제별 기계적 품질계수를 반영한 온라인자료의 집서수준의 수정

주제	총류	철학	종교	사회과학	자연과학	기술과학	예술	언어	문학	역사
현행집서수준(A)	3	3	3	3	3	3	3	3	3	3
기계적품질계수(B)	0.73	0.72	0.31	1.66	0.28	1.53	0.82	0.45	1.15	1.75
수정집서수준(C)	2.19	2.16	0.93	4.98	0.84	4.59	2.46	1.35	3.45	5.25
최종집서수준(D)	2	2	1	5	1	5	2	1	3	5

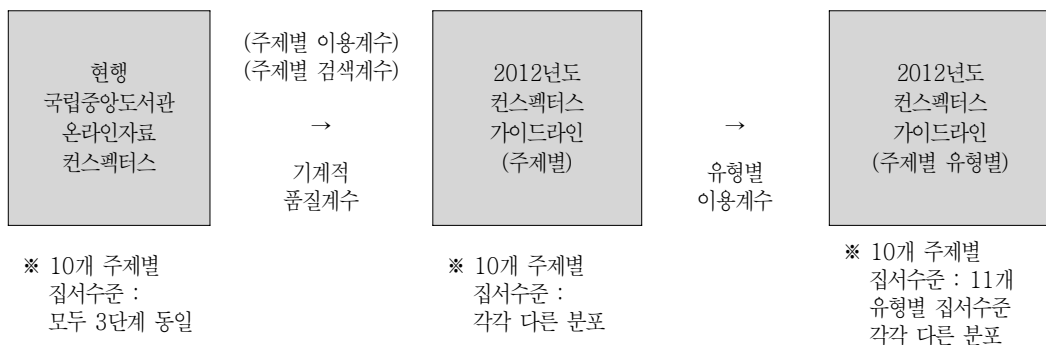
위에서 도출된 새로운 주제별 집서수준은 해당 주제를 대표하지만 온라인자료의 특성상 같은 주제 내에서도 유형에 따라 이용율이 큰 차이가 있으므로 이를 모든 유형에 동일하게 적용하는 것보다는 유형별 이용계수를 주제별 집서수준에 곱하여 유형별 집서수준을 구별하였다. 이렇게 하여

주제별, 유형별로 기계적 품질을 안배한 새로운 온라인자료 컨스펙터스 운영모형을 제시하면 <표 12> 및 <그림 1>과 같다.

<표 12> 주제별, 유형별 기계적 품질을 반영한 온라인자료 컨스펙터스 운영모형

유형별 (이용계수)	주제별									
	총류	철학	종교	사회 과학	자연 과학	기술 과학	예술	언어	문학	역사
최종집서수준	2	2	1	5	1	5	2	1	3	5
수정집서수준	2.19	2.16	0.93	4.98	0.84	4.59	2.46	1.35	3.45	5.25
전자책(0.58)	1.27	1.25	0.53	2.88	0.48	2.66	1.42	0.78	2.00	3.04
전자저널(1.55)	3.39	3.34	1.44	7.71	1.30	7.11	3.81	2.09	5.34	8.13
동영상자료(1.00)	2.19	2.16	0.93	4.98	0.84	4.59	2.46	1.35	3.45	5.25
이미지(0.42)	0.91	0.90	0.39	2.09	0.35	1.92	1.03	0.56	1.44	2.20
사운드(1.00)	2.19	2.16	0.93	4.98	0.84	4.59	2.46	1.35	3.45	5.25
소프트웨어(0)	0	0	0	0	0	0	0	0	0	0
데이터세트(0)	0	0	0	0	0	0	0	0	0	0
콜렉션(0)	0	0	0	0	0	0	0	0	0	0
웹사이트(0.18)	0.39	0.38	0.16	0.89	0.15	0.82	0.44	0.24	0.62	0.94
장애인대체자료(6.50)	14.23	14.04	6.04	32.37	5.46	29.83	15.99	8.77	22.42	34.12
기록물(0.47)	1.02	1.01	0.43	2.34	0.39	2.15	1.15	0.63	1.62	2.46

※ 최종집서수준은 현행집서수준과의 비교를 위하여 편의상 수정집서수준의 소수점이하를 반올림한 것이며, 이하 전자책부터 기록물까지의 집서수준은 동일한 주제 내에서 유형별 이용계수의 차이를 반영한 주제내 유형간 이용가치의 경중을 비교하기 위한 것임.



<그림 1> 국립중앙도서관 온라인자료 컨스펙터스 운영모형 개념도

V. 결 론

디지털도서관은 종이책 중심의 전통적 도서관과는 달리 망라적 수집과 이용이 쉽지 않다. 특히 빅데이터시대를 맞아 모든 온라인자료를 파악, 수집, 관리, 이용하는 것은 현실적으로 불가능하다. 그리고 온라인자료는 매체의 유형이 다양하고 끊임없이 진화하고 있기 때문에 이러한 온라인자료의 특성을 감안한 장서개발정책이 수립될 필요가 있다. 도서관의 주제별 자료수집 의지와 수준을 나타내는 컨스펙터스도 온라인자료에 맞는 방식으로 운영할 필요가 있다.

국립중앙도서관은 2009년 9월 26일 개정·시행된 도서관법 제20조의2에 의하여 대한민국에서 서비스되는 웹사이트, 웹문서와 같은 공개용 온라인 자료뿐만 아니라 전자책, 전자저널과 같은 판매용 온라인 자료 중 보존가치가 높은 온라인자료의 수집·보존을 위해 전자출판물 중 전자책과 학회지를 우선 수집대상 자료로 선정하여 중점수집하고 있으며, 2011년의 경우 약 17억원의 예산으로 전자책 7,593책, 학회지 909종 160,569건, 이미지 9종 37,122건, 오디오북 59종 643건, 이러닝 자료 15종 215건을 수집하였다. 그리고 공개용 온라인자료에 있어서도 2011년도에 웹문서 128,822건, 웹사이트 5,012건, 총 133,834건 등 718,399건의 공개용 온라인 자료를 수집하였다.

그런데 비록 온라인자료 수집에 대한 법규정이 보존가치가 높은 자료를 선정, 수집, 보존하도록 명시하고 있다 하더라도 국립중앙도서관 디지털도서관의 온라인자료 수집에 있어서 이용측면을 함께 고려하는 것이 바람직하다. 즉, 디지털도서관이 온라인자료를 확충할 때 모든 주제와 유형의 자료를 동일한 수준으로 수집하는 것 보다는 자료의 활용도와 이용자의 만족도를 높일 수 있는 기계적 품질이 높은 주제의 자료 수집을 확대하는 것이 필요하다. 이에 자료의 활용도를 나타내는 이용율과 검색율에 근거한 기계적 자료품질계수를 주제별로 산정하고, 여기에 현행집서수준을 곱하여 최종적인 주제별 수정집서수준을 도출한 온라인자료 컨스펙터스 운영모형을 제안하였다.

그 결과 이용율이 높은 기술과학과 검색율이 높은 역사와 사회과학 등 주제의 집서수준을 대폭 상향조정하고 여타 주제들은 현행유지 또는 하향조정 되어야 하는 것으로 나타났다. 그리고 동일한 주제 내에서도 다양한 매체유형별 기계적 품질계수를 감안한 안배가 필요한 것으로 나타났다.

다만 도서관법에서 명시하고 있는 보존가치가 높은 자료의 보존 필요를 위해서 위 모형상 집서수준이 현행보다 하향조정되는 주제에 있어서도 보존가치가 높은 자료에 대하여는 현행집서수준을 그대로 적용할 필요가 있다. 이 경우 현행보다 하향조정되는 모든 주제에 현행수준인 3단계 교육학 습수준을 부여하는 방법과 주제별 최종집서수준은 모형대로 하향조정하되 주제 내에서 보존가치가 높은 자료만 현행수준을 부여하는 방법이 있으나 이는 실제 운영과정에서 선택할 수 있을 것이다.

아울러 현행법이 온라인자료의 보존에 대하여 언급하고 있는 것은 이용에 대한 중요성을 간과하려는 것이 아니라 보존을 강조하는 취지라고 보아야 하므로 보다 분명한 온라인자료 수집방향의

설정을 위하여 보존과 이용을 함께 강조하는 방향으로의 개정을 검토할 필요가 있다고 생각된다.

이러한 컨스펙터스 운영모형은 기존의 컨스펙터스의 틀과 목표를 완전히 수정하려는 것이 아니라 매년 온라인자료의 기계적 품질계수를 측정하여 당해 연도의 컨스펙터스 운영모형(연도별 가이드라인)으로 활용하자는 것이다. 이러한 측면에서 이 모형은 수정이 아니라 운영모형인 것이다.

온라인자료의 컨스펙터스를 이 모형대로 운영할 경우 국립중앙도서관의 경우 연간 20억에 가까운 온라인자료 수집의 선택과 집중이 이루어져 예산집행의 효율성이 더욱 증대될 것으로 기대된다. 그리고 이 운영모형은 여타 디지털도서관의 온라인자료 수집 및 이용모델로 활용될 수 있을 것이다.

본 연구의 한계 및 후속연구와 관련하여 먼저 온라인자료의 유형 중 전자책, 전자저널, 이미지, 기록물 유형은 디지털원문을 포함하고 있는데, 디지털원문은 이용계수와 100대 검색어 관련자료 소장비율이 높으므로 이에 상응하는 디지털원문의 기계적 품질계수를 별도로 산정하여 가중치를 부여할 필요가 있으나 이를 위해서는 온라인자료일반과 디지털원문자료 사이의 통계적 호환성이나 등간성에 대한 추가적인 연구가 필요하여 본 연구에서는 그 필요성만을 언급하고자 한다.

그리고 위에서 제안한 운영모형이 보다 더 높은 수준의 타당성을 갖기 위해서는 기계적 품질 외에도 인간적 품질에 대한 평가가 함께 반영될 필요가 있다. 후속연구에서는 이러한 새로운 컨스펙터스 운영모형을 검증하고 타당성을 높이기 위하여 이용자들을 대상으로 주제별 중요도와 만족도 등에 대한 설문조사를 통해 온라인자료의 인간적 품질을 측정하여 운영모형에 반영한다면 더 좋은 결과를 도출할 수 있을 것이다.

참고문헌

- 박재용. “공공도서관의 이용자만족도에 관한 한미간 비교사례연구.” 한국비블리아학회지, 제23권, 제2호(2012. 6), pp.151-169.
- 박진희. “컨스펙터스방법을 이용한 장서평가연구.” 정보관리학회지, 제15권, 제2호(1998. 12), pp.164-191.
- 윤희윤, 김일영. “이용계수를 적용한 공공도서관 소설장서 평가.” 정보관리연구, 제42권, 제4호(2011. 12), pp.175-194.
- 이진경, 이지연. “이용자 장서이용행태 분석을 통한 공공도서관의 장서개발에 관한 연구.” 한국비블리아학회지, 제22권, 제3호(2011. 9), pp.151-171.
- 이정수, 남영준. “행위이론을 적용한 도서관 이용자 컨텍스트 정보의 CAbAT 모델링.” 한국도서관정보학회지, 제43권, 제1호(2012. 3), pp.1-19.
- 이혜원. “목록의 집중기능을 향상시키는 원형개념에 관한 연구.” 한국비블리아학회지, 제22권, 제3

호(2011. 9), pp.91-107.

장덕현. “전자자원선정을 위한 컨스펙터스 수정모형연구.” 한국비블리아학회지, 제20권, 제2호 (2009. 6), pp.31-44.

장덕현, 서태설. “디지털 장서개발정책 기본요소 연구.” 한국도서관정보학회지, 제40권, 제3호 (2009. 9), pp.1-21.

정은경. “멀티미디어 정보요구와 검색행태에 관한 탐색적 연구.” 한국문헌정보학회지, 제46권, 제3호(2012. 9), pp.251-270.

