

전자저널 관리를 위한 이용통계의 효과적 활용 방안*

Toward the Effective Utilization of Usage Statistics for the Management of Electronic Journals

김 성 진**
Sung-Jin Kim

차 례

1. 서론	4. 전자저널 관리를 위한 COUNTER 이용통계 활용 방안
2. 관련 연구 동향	5. 결론 및 제언
3. 전자저널의 COUNTER 이용통계 제공 현황	· 참고문헌

초 록

제한된 예산, 고가의 패키지 구독비, 출판사 주도의 라이선스 협상 등의 불리한 환경에서 자관의 실정에 맞는 전자저널을 구성하기 위해 도서관은 전자저널 이용데이터를 수집하고 분석해야 한다. 본 연구는 전자저널 이용통계 연구 동향 및 COUNTER 3판에 근거한 이용통계 제공 현황을 살펴보고 현장에서 쉽게 활용 가능한 이용통계 분석의 가이드라인을 제시하고자 하였다. 현재 제공되는 이용통계 보고서를 기반으로 분석할 수 있는 핵심 이용통계치로 저널별 원문이용도, 저널이용률, 원문이용단가, 고이용군, 저이용군을 제안함으로써 전자저널 관리를 위한 이용통계의 효과적인 활용 방안을 다루었다.

키 워 드

전자저널, 이용통계, 이용데이터, COUNTER, 저널보고서

* 이 논문은 2009학년도 인하공업전문대학 교내연구비 지원에 의하여 연구되었음.
 ** 인하공업전문대학 비서과 조교수
 (Assistant Professor, Dept. of Secretary Service, Inha Technical College, sjkim@inhac.ac.kr)
 • 논문접수일자: 2010년 8월 19일
 • 최종심사일자: 2010년 9월 27일
 • 게재확정일자: 2010년 10월 7일

ABSTRACT

Libraries are encountering hostile conditions around journal licensing such as limited budget, the high price of packages, and vendor-led negotiation. They need to collect and analyze usage data of electronic journals to develop electronic journal collection appropriate for their own circumstances. The purpose of this study is to suggest an practical guideline for librarians' analysis of electronic journal usage statistics. For this, the study reviewed related previous studies and examined current state on usage statistics provided from COUNTER release 3 compliant vendors. Finally this study proposed five core statistics including full-text article request per journal, journal using rate, price per full-text article request, most use group, and low use group, and further discussed how to use them effectively for the electronic journal management.

KEYWORDS

Electronic Journal, Usage Statistics, Usage Data, Use Data, COUNTER, Journal Report

1. 서론

1.1 연구의 필요성

자관의 이용 및 요구에 따라 도서관이 독자적으로 장서를 선별, 수서하던 과거와 달리 현재 전자저널 구독의 대부분은 벤더 및 출판사로부터 패키지 형태로 제공되고 있다. 구독하고자 하는 해당저널 이외에 별도의 저널들이 함께 구성되는 등 도서관이 전적으로 주도하던 장서개발 영역이 과거보다 상대적으로 많이 축소되었다. 전자저널 도입 초기에 전자저널이라는 생소한 자원을 다루는 것이기에 도서관은 이를 제공하는 업체 측에 수동적 또는 소극적으로 대응할 수밖에 없었던 것이

사실이다.

제한된 예산, 고가의 패키지 구독 비용, 제 공업자 주도의 라이선스 협상 등의 불리한 환경이 여전히 지속되고는 있지만 도서관도 이제 전자저널이라는 정보자원을 다루는 데 있어서 다소 안정을 찾았을 법한 시간이 흘렀다. 도서관이 보다 적극적으로 자관의 실정에 맞는 전자저널을 구성하고 관리해 나가기 위해서는 전자저널에 대한 자관 이용자들의 요구와 이용을 파악하는 것이 무엇보다 중요해졌다.

과거 인쇄환경에서 도서관의 이용률을 파악하는 주요 지표로 도서관 방문 수와 대출기록이 사용되었지만, 대출이 제한된 저널의 경우에는 재서가 통계법을 사용한다 할지라도 그 이용률을 정확하게 파악하는 것이 어려웠다.

반면에 전자저널은 서버에 기록되는 로그 데이터를 통해 자관이 구독하는 전자저널에 대한 이용 현황을 비교적 용이하게 분석할 수 있다는 장점을 가진다.

신뢰할 수 있는 정확한 이용통계는 계약협상, 예산지출의 정당화, 비용효과 분석 등에 활용되어 도서관의 의사결정에 도움을 줄 수 있다. 또한 전자저널에 대한 자관 이용자의 요구와 행태를 이해할 수 있게 해주며, 이용자 교육, 홍보 등의 전략적 계획을 수립하는 데 도움을 준다(Shim & McClure 2002). 전자저널이 등장한 지 10여년이 흐른 지금, 전자저널 이용데이터의 잠재적 능력에 대한 기대와 함께 이용통계의 이용가능성, 활용 전략, 표준안, 가이드라인 개발 등의 활동이 활발하다. 이러한 노력 끝에 COUNTER(Counting Online Usage of NeTworked Electronic Resources)와 같은 표준화 지침이 지속적으로 개발되고 보완되고 있으며, 전자저널 이용통계를 제공하는 벤더의 수가 증가하고 있다.

하지만 이러한 발전적인 추세 속에서도 전자저널 이용통계의 활용은 여전히 어려운 수준에 머물러 있다. 우선적으로 도서관이 구독하는 전자저널의 범위가 증가하였기 때문에 도서관 사서가 수집해야 하는 이용통계의 양이 많으며, 각기 다른 여러 벤더로부터 통계 데이터를 수집하는 데에도 상당한 시간과 노력이 든다. 또한 패키지나 서비스별로 각각 제공된 이용통계 보고서를 통합하는 데 많은 에너지를 소비해야 하고 이를 분석하는 데 필요한 시간

과 지식이 충분하지 않는 것도 이용통계의 활용을 막는 장애요인으로 작용하고 있다.

1.2 연구의 목적

본 연구는 전자저널 이용데이터를 어떻게 분석하며, 어떠한 이용통계치를 전자저널 관리 및 의사결정 도구에 활용할 수 있는지 그 실용적인 방안을 제시하고자 한다. 전자저널 이용통계 분석에 대한 많은 선행연구들이 이루어졌지만 대부분 복잡한 통계치를 제시한 경우가 많다. 특정 목적에 의해 전자저널 이용행태에 대한 심도있는 분석이 필요한 경우에는 이러한 연구 및 분석법이 의미가 있으나 일상적인 도서관 사서의 책무로서 실무에 실제적으로 활용되기에는 무리가 있다. 사실상 이용통계의 중요성을 인지하나 막상 이용통계 데이터에 접한 사서들은 이를 어떻게 분석하고 활용해야 하는지 난감해하기 일수이다.

이에 본 연구는 실무에 적합한 핵심적인 이용통계치를 추출하고 이를 분석 및 활용하는 방안을 제시함으로써 현장의 사서들로 하여금 전자저널 관리를 위해 이용통계를 쉽게 활용하고 효과적으로 적용할 수 있도록 하였다. 이를 위해 전자저널 이용통계 관련 연구 동향을 살펴보았으며, COUNTER 3판을 기준으로 전자저널 이용통계의 제공 현황을 분석하였다. 이용통계 분야는 학계에서의 분석적인 연구도 물론 중요하지만 현장에서의 실제적인 활용과 지속적인 실무가 보다 중요하다는 점에서 본

연구는 현장의 사서들이 이용통계 연구의 복잡성을 이해하는 데 도움이 될 것으로 본다.

2. 관련 연구 동향

전자저널 이용통계 연구는 2000년 이후부터 증가하기 시작하였으며 크게 이용통계의 제공 및 활용 전략에 관한 연구, 이용통계의 표준화에 관한 연구, 이용통계에 기반을 둔 이용패턴 분석 연구로 나눌 수 있다.

2.1 이용통계의 제공 및 활용

이용통계의 제공 및 활용에 관한 대표적인 연구로, Luther(2001)는 도서관 및 출판사와의 인터뷰를 통해 전자저널 이용통계와 관련하여 양측에 미칠 주요 이슈들을 포괄적으로 다루었다. 이 연구는 측정될 이용통계의 데이터 요소, 데이터의 신뢰성을 위한 고려사항, 도서관 및 출판사의 향후 역할에 대해 논의함으로써 전자저널 이용통계에 관한 큰 그림을 제시한 초기연구로 의의를 갖는다.

Hahn & Faukner(2002)는 전자저널 이용통계에서 가장 중요한 요소를 ‘논문 수’로 보고 이를 기반으로 전자저널 선별에 유용한 3개의 이용 통계치를 제시하였다. 구독비용을 이용된 논문 수로 나누는 ‘이용된 논문의 평균 단가’(average cost per access), 구독비용을 온라인에서 제공하는 전체 논문 수로 나누는

‘전체 논문의 평균 단가’(average cost per article), 논문의 다운로드 횟수를 전체 논문 수로 나누는 ‘수록된 논문 기준 이용률’(content-adjusted usage)을 근거로 구독 중인 전자저널을 평가하고 잠재적인 구매대상 저널을 평가하는 방법을 선보였다. 이들이 제시한 통계치는 향후 여러 논문(박은경 2005; 황옥경 2007 등)에서 응용된 바 있으나 ‘논문 수’가 쉽게 측정가능하지 않다는 점에서 실무 활용에 어려움이 따른다.

Baker & Read(2008)는 미국 및 캐나다 연구도서관을 대상으로 벤더가 제공하는 전자저널 이용데이터의 활용에 관한 설문조사를 실시하였다. 벤더가 제공한 이용통계를 활용하는 데 어느 정도의 노력을 투자하고 있는지, 데이터가 어떻게 이용되는지, 어떤 데이터가 사서들에게 가장 유용한지를 알아보려고 하였다. 이용데이터의 표준화 부족과 불일치성이 효율적인 이용통계 활용의 장애물로 꼽혔고, 이용데이터를 분석하기보다 사실상 수집하고 재편집하는 데 더 많은 시간을 할애하고 있었으며, ‘원문 다운로드 수(number of full-text downloads)’가 가장 유용한 통계치로 평가되었다.

전자저널 이용통계 제공 및 활용에 관한 국내 연구로, 김혜선(2004)은 한국과학기술정보연구원에서 구독 중인 전자저널 패키지 중에서 비교적 충실하게 통계를 제공하는 8개의 DB를 선정하여 이들의 통계서비스를 평가하였다. COUNTER 표준지침 발표 이후 전자저널 이용통계 서비스가 많이 개선되었으나 장

서평가에 활용되기에는 여전히 미흡하며, 향후 도서관, 관련 단체, 출판사 등의 협력을 통해 보다 구체적이고 표준화된 이용통계 요소가 개발되어야 함을 논하였다. 심원식(2005)은 전자정보의 이용통계에 대해 이해하고 도서관 실무에서 뿐만 아니라 학계에서 이용통계가 활용될 수 있도록 데이터 수집, 처리, 분석 등의 제반 사항에 관한 활용 전략을 제시하였다. 전자정보 관련 의사결정 및 업적 개선을 위해 이용통계를 사용하고자 할 때 개별 도서관이 고려해야 할 사항 및 활용 방안이 구체적으로 설명되고 있어 전자정보 이용통계 활용의 좋은 가이드라인 역할을 하고 있다.

황옥경(2007)은 서울 및 경기 소재 48개 대학도서관을 대상으로 국내 대학도서관에서의 전자저널 이용통계 제공 및 활용 현황을 분석하였다. 특히 선행연구 분석을 통해 전자저널 평가에 활용될 수 있는 5가지 통계치를 '1회 이용 단가', '논문 1편당 이용 횟수', '이용된 논문 수만을 기준으로 계산된 논문 1편당 이용 횟수', '전체 논문 중에서 이용된 논문의 비율을 알아보는 충실도', '논문 1편의 단가'로 제시하고 이 통계치들의 활용 현황을 조사하였다. 주목할 만한 조사 결과로는 현재 제공되는 이용통계 데이터에 대한 만족도가 낮고, 출판사별로 상이한 데이터를 제공하고 있어 비교 가능한 데이터가 부재하며, 구독 대학도서관 전체의 평균 이용률에 대한 정보가 제공되지 않는 등 이용통계를 분석할 기준점이 모호하다는 점에서 높은 불만을 보인 점이다.

2.2 이용통계의 표준화

이용통계의 표준화에 관한 연구는 표준화 방안 및 표준화 프로젝트 활동을 다루거나, 이용통계 데이터의 신뢰성 및 타당성에 관한 분석이 주로 이루어졌다. Shim & McClure(2002)는 전자정보 서비스의 통계 및 평가방법 구축을 목표로 하는 ARL E-Metrics 프로젝트의 주요 활동 및 결과에 대해 보고하였다. 도서관의 의사결정에 유용한 보고서의 수준을 평가하기 위해 주요 8개 벤더의 이용통계를 테스트한 결과, 전반적으로 벤더들의 이용통계는 쉽게 수집 및 처리 가능했으나 핵심 이용통계치 및 보고서 포맷의 표준화, 제공된 통계치의 타당성 검증은 여전히 중요한 이슈로 남아있다고 주장하였다. Kraemer(2006)는 이용데이터의 무결성을 강조하며, 일관되고 정확한 이용통계의 수집 및 통합을 위해 고려해야 할 사항들을 다루었다. 특히 링킹 옵션, Open-URL 등과 같이 전자저널로 이끄는 접근체계의 유형에 따른 영향과 신생 저널이 전자저널 이용에 미치는 영향을 논함으로써 이용통계 분석 및 해석에 주의할 점을 지적하였다.

Pesch(2006)는 COUNTER 프로젝트의 배경을 소개하였을 뿐만 아니라 메타검색(federated searching), 프리페칭(pre-fetching) 등의 기술이 이용통계 보고에 미칠 부정적인 영향을 검토하였으며, Blecic, Fiscella, & Wiberley(2007)는 벤더가 제공하는 통계와 COUNTER 표준을 따르는 통계간의 차이를 분석

하고 정보주지서비스, 메타검색 등의 혁신적 기능이 이용통계 요소인 세션과 검색의 의미를 어떻게 변화시키는지를 논하였다.

국내의 경우, 정은경(2008)은 전자저널 이용통계의 개념과 표준화 필요성 및 관련 프로젝트를 살펴보았으며, 특히 국내의 전자저널 제공기관 4곳을 대상으로 설문조사를 실시함으로써 국내 전자저널 이용통계의 표준화 현황 및 요구에 대한 현황을 파악하였다. 이를 통해 국내 출판사 및 대형업체의 이용통계 제공 및 표준화 노력을 분석하고 전자저널 이용통계 표준화 방안을 제시하였다.

2.3 이용통계에 기반한 전자저널 이용 분석

이용통계를 활용하여 전자저널의 이용 패턴을 분석한 연구 중에서 Conyers & Dalton(2005)은 전자저널 이용에 대한 국가적인 수준의 개략을 파악하기 위해 다양한 규모의 대표 기관들을 선정하여 전자저널 이용데이터를 심층 분석하였다. NESLi2에 참여한 4개 출판사가 2004년 1월~6월 이용데이터를 제공하였고, 2003년 동일기간의 데이터도 비교를 위해 사용되었다. 이용현황 분석을 위해 원문 요청 수, 이용대비 단가, 높게 혹은 낮게 이용된 저널 분석, 주제별 이용, 인쇄본과 병행구독된 타이틀에 대한 분석, FTE별 단가 등의 통계치를 이용하였으며 이용통계에 영향을 미치는 요인들을 분석하고, 이를 바탕으로 JISC를 위한 이용통계 관련 권고안을 제시하였다.

Evans & Peters(2005)는 Emerald에서 제공하는 경영 및 관리 분야의 전자저널 패키지인 Emerald Management Xtra(EMX)를 대상으로 과거 인쇄저널 환경에 적용되었던 80:20 법칙이 전자저널의 경우에도 적용되는지를 실험하고자 하였다. 조사 결과, 47.4%의 저널이 전체 이용의 80%를 차지함을 밝힘으로써 과거의 파레토 법칙을 많이 초과하는 이용 패턴을 보였음을 밝혔다.

국내의 경우, 박은경(2005)은 이화여대 중앙도서관이 가입한 전자저널 컨소시엄의 2002년~2004년 이용데이터를 수집하여 이용빈도, 이용효율성, 이용일치성, 이용안정도, 비용효율성, 원문 매체별 선호도, IP 분석을 실시하였다. 전자저널 이용 행태를 파악하였을 뿐만 아니라 이용통계 분석 자료를 바탕으로 전자저널 구독 형태를 유형화하고 유형별 수서정책 방안을 제시하였다는 점에서 의의가 있다.

이 외에도 전자저널 이용통계치를 전자저널 평가지표로 대체 활용될 수 있는지에 대한 연구, 이용통계 전송에 관한 기술적 연구 등과 같은 연구가 진행 중에 있다. 그러나 위에서 살펴본 바와 같이 국외 연구에 비해 전자저널 이용통계에 관한 국내 연구의 수가 상대적으로 많이 부족하다. 전자저널에 관한 국내 연구는 가격모형 및 아카이빙에 보다 초점이 맞추어져 있는 것이 현실이며 현장 사서 및 학계에서도 이용통계의 중요성은 인식하나 아직 이용통계의 수집 및 활용은 활발히 이루어지지 못하고 있다. 따라서 전자저널 이용통계에서 제공되는 데이터 요소에 대

한 이해를 도모하고 전자저널 관리 및 의사결정에 적용될 수 있는 핵심 통계치를 파악하고 그 활용 전략에 대한 연구가 필요한 실정이다.

3. 전자저널의 COUNTER 이용통계 제공 현황

3.1 COUNTER 3판의 이용통계 보고서 유형

COUNTER는 여러 벤더들로부터 이용통계가 제공되더라도 서로 비교 가능한 이용통계가 될 수 있도록 하기 위한 목적에서 온라인 이용을 측정하는 방안의 표준화를 도모하였다(Pesch 2006). 이용통계와 관련해 ICOLC(International Coalition of Library Consortia), PIRUS(Publisher and Institutional Repository Usage Statistics), E-Metrics 등의 다양한 프로젝트가 존재하지만, COUN-

TER는 가장 활발한 움직임을 보이며 전 세계적으로 많이 채택되고 있는 지침으로써 현재 전자저널 이용통계의 표준화를 주도해가고 있다.

COUNTER는 저널과 DB에 초점을 맞춘 시행령 1판을 2002년 12월에 발표한 이후에 2005년 4월에 시행령 2판을, 2008년 8월에 시행령 3판을 발행하였다. 1판에 비해 2판은 저널보고서 1을 확장한 것이 특징이었고, 3판은 SUSHI(Standardized Usage Statistics Harvesting Initiative) 프로토콜의 사용을 의무화하고 XML 포맷의 보고서 제공을 요구하며, 36개의 로봇 및 웹크롤러 목록을 제공하여 이용자의 실제 이용이 아닌 인위적인 접근을 통계에서 배제하도록 함으로써 이용통계의 수집을 용이하게 할 뿐만 아니라 이용통계 데이터의 신뢰성을 높이고자 하였다(COUNTER 2008b).

COUNTER 3판에서 규정하는 이용통계 보고서의 유형은 총 12가지로, 각 보고서 유형에 대한 간단한 설명은 <표 1>과 같다(CO-

<표 1> COUNTER 3판의 이용통계 보고서 유형

분류	보고서명	약칭	설명	제공형식
저널	Journal Report 1	JR1	저널별/월별 원문요청 건수	엑셀, CSV, XML
	Journal Report 1a	JR1a	저널별/월별 아카이브로부터의 원문요청 건수	엑셀, CSV, XML
	Journal Report 2	JR2	저널별/월별 접속거절 수	엑셀, CSV, XML
	Journal Report 5	JR5	저널별/출판년도별 원문요청 건수	엑셀, CSV, XML
DB	Database Report 1	DB1	DB별/월별 검색건수와 접속건수	엑셀, CSV, XML
	Database Report 2	DB2	DB별/월별 접속거절 수	엑셀, CSV, XML
	Database Report 3	DB3	서비스별/월별 검색건수와 접속건수	엑셀, CSV, XML
컨소시엄	Consortium Report 1	CR1	월별 저널 논문 및 전자책 캡처에 대한 요청 건수	XML
	Consortium Report 2	CR2	DB별/월별 검색건수	XML
추가 옵션	Journal/Book Report 1	JB1	타이틀별/월별 아이템 요청 건수	XML
	Journal Report 3	JR3	저널별/월별/페이지타입별 아이템 요청 건수	엑셀, CSV, XML
	Journal Report 4	JR4	서비스별/월별 검색건수	엑셀, CSV, XML

UNTER 2008c). COUNTER 3판에는 4개의 저널 보고서, 3개의 DB 보고서, 2개의 컨소시엄 보고서가 표준 형식으로 제공되며 그 외에 3개의 옵션 보고서가 추가적으로 제시되어 있다. 각 보고서에 들어가는 이용통계 항목을 살펴보면 원문요청 건수, 접속거절 수, 검색건수, 접속건수의 4가지 데이터 요소가 기반이 되어 저널별, DB별, 서비스별, 월별, 출판년도별 등의 이용통계 보고서가 제공되고 있다.

COUNTER 실행지침은 각 유형별 보고서의 형식을 <그림 1>과 같이 제시하고 있다. 예를 들어, JR1은 셀 A1에 보고서 유형을 명시하고 A2에는 보고서의 기준이 되는 기관명 또는 컨소시엄명을 적어야 하며 A6에는 'Total for all journals'라고 적고 F6부터는 월별 원문요청 건수 합계와 당해연도(Year to Date) 합계를 제시해야 한다. 이처럼 스프레드시트의 셀 위치를 정확히 준수하여 보고서를 제공하도록 규정하고 있기 때문에 여러 벤더로부터 제공받은 이용통계 데이터를 파악하고 활용하는 것이 보다 수월해지게 되었다.

COUNTER 3판에서 새로 제안된 JR1a는

아카이브에 대한 원문 이용을 파악하기 위한 것으로 아카이브를 제공하는 벤더는 JR1a나 저널 출판년도별 원문요청 건수를 기록한 JR5의 둘 중 하나를 반드시 제공해야 한다. 또한 JB1은 전자저널과 전자책을 같은 플랫폼에서 제공하는 벤더가 사용할 수 있도록 제안된 것이며, JR3와 JR4는 저널별/월별 원문요청 건수만을 제공하는 JR1의 내용을 보충하기 위해 추가적으로 페이지 타입별 요청건수와 검색건수를 각각 제공할 수 있는 보고서 형식이다 (COUNTER 2008a).

COUNTER의 특징 중 하나는 원문요청 건수, 접속거절 수, 검색건수, 접속건수와 같은 4개의 데이터 요소만을 통해 비교적 간단하지만 기본적인 이용통계 보고서의 제공을 권장하는 점이다. 일부에서는 데이터 요소가 너무 단순하여 이용통계가 의미하는 바가 불충분하지 않을까 하는 우려의 목소리를 내기도 한다. 그러나 제공되는 데이터 요소가 많아지면 여러 벤더로부터 받은 이용통계를 분석해야 하는 담당 사서의 업무가 복잡해질 뿐만 아니라, 많은 데이터 요소를 집계해야 하는 벤더 측의

	A	B	C	D	E	F	G	H	I	J	K
1	Journal Report 1 (R3)	Number of Successful Full-Text Article Requests by Month and Journal									
2	<Criteria>										
3	Date run:										
4	yyyy-mm-dd										
5		Publisher	Platform	Print ISSN	Online ISSN	Jan-2009	Feb-2009	Mar-2009	YTD Total	YTD HTML	YTD PDF
6	Total for all journals		Platform Z			772	972	1165	2909	1032	1877
7	Journal of AA	Publisher X	Platform Z	1212-3131	3225-3123	456	521	665	1642	522	1120
8	Journal of BB	Publisher X	Platform Z	9821-3361	2312-8751	203	251	275	729	290	439
9	Journal of CC	Publisher Y	Platform Z	2464-2121	0154-1521	0	0	0	0	0	0
10	Journal of DD	Publisher Y	Platform Z	5355-5444	0165-5542	113	200	225	538	220	318

출처: COUNTER(2008c), p. 10.

<그림 1> COUNTER 3판의 JR1 형식 예

부담도 증가하게 된다. Schufreider & Romaine(2008)도 현재 COUNTER 표준안이 단 순해 보일지 모르나 지금은 도서관, 출판사, 벤더가 함께 모여 공통안을 도출하는 단계이 므로 값진 데이터를 제공할 수 있는 개선된 표 준안을 구축하는 데 초점을 맞추는 것이 보다 바람직하다고 언급한 바 있다.

또 다른 특징으로는 COUNTER 표준은 이 용데이터가 언제 어떻게 측정되어야 하는지에 대한 가이드라인을 잘 제시하고 있다는 점이다. 예를 들어, 원문요청 건수를 계산하는 데 에 있어서 다음과 같은 기준을 적용시켜야 한다. 이용자가 느린 접속 환경으로 인해 PDF 원문보기 버튼을 여러 번 클릭했을 경우 이를 여러 번의 요청으로 보면 안 된다. HTML 문 서는 10초 이내에, PDF는 30초 이내에 같은 논문을 2번 이상 클릭하였을 경우 이를 모두 한 번으로 계산하도록 조정하고 이때에는 마 지막에 이루어진 클릭 시간을 기준으로 원문 요청 건수를 계산해야 한다. 그리고 접속건수 와 검색건수의 측정을 위하여 세션 시간을 30 분으로 정의하였고 검색결과가 0건인 경우에도 검색으로 기록하도록 지시하고 있다. 진정한 검색에 대한 이용통계 추출을 위해서 CO-UNTER는 하나의 질의어가 시스템을 통해 여

러 DB로 보내지는 메타검색 및 자동검색에 의 해 생성된 접속건수와 검색건수를 별도로 분 리하여 제공할 것을 명시하고 있다. 또한 인터넷 로봇 및 웹 크롤러, 그리고 LOCKSS와 같은 보존을 위한 캐시에 의해 생성된 데이터 는 이용통계 보고서에서 배제할 것을 규정하 고 있다.

3.2 이용통계 보고서 제공 벤더 현황

2010년 8월에 제공된 COUNTER의 벤더 목록¹⁾을 살펴봄으로써 COUNTER 3판을 준 수하는 벤더의 현황을 살펴보고자 한다. 비교 를 위해 COUNTER 2판과 1판의 벤더 목록²⁾ 을 함께 참조하였다.

COUNTER 3판을 준수하는 벤더는 총 113 곳으로, COUNTER 2판을 준수했던 벤더가 2007년 7월 당시 66곳이었던 것과 비교하였 을 때 그 수가 약 72% 증가하였다. COUNTER 1판을 준수했던 벤더 수가 2005년 12월 당시 14곳으로 시작했던 것을 감안하면 COUNTER 표준을 따르는 벤더의 수가 큰 폭으로 증가하 였으며 그만큼 표준화된 이용통계 보고서의 제 공이 안정화되어 가고 있음을 짐작할 수 있다.

그러나 COUNTER 1판에서 2판으로의 작

1) "Register of Vendors." [인용 2010. 8. 9]. <<http://www.projectcounter.org/compliantvendors.html>>.

2) 2판을 준수하는 2007년의 벤더 목록은 <<http://www.projectcounter.org/articles.html>>에서 참조하였으며([인용 2010. 8. 9]), 1판을 준수하는 2005년의 벤더 목록은 미국의 예일대학교도서관을 중심으로 운영되는 메일링 리스트 인 LIBLICENSE-L의 아카이브 자료 중 "Update on usage statistics for informaworld"에 대한 토론(<<http://www.library.yale.edu/~llicense/ListArchives/0703/msg00060.html>>)에서 참고하였다([인용 2010. 8. 11]). LIBLICENSE-L은 대학 및 연구도서관의 전자콘텐츠 라이선스와 관련된 주제에 대해 토론하기 위해 구축된 메일 링 리스트이다.

은 변화와 달리, 3판에서는 각기 다른 벤더들로부터 제공되는 이용통계 보고서의 통합이 손쉽게 이루어질 수 있도록 SUSHI 프로토콜을 사용할 것과 XML 형식의 보고서를 제공할 것을 의무화하였을 뿐만 아니라 제공하는 보고서의 유형을 확장함으로써 벤더들의 부담을 가중시킨 것이 사실이다. COUNTER 2판을 준수하던 66개의 벤더들 중에서 현재 COUNTER 3판을 준수하는 벤더 목록에 포함된 기관을 조사해본 결과, 절반에도 미치지 못하는 29개의 벤더들만 COUNTER 3판으로의 이행을 성공한 것으로 나타났다.

COUNTER는 벤더들이 COUNTER 표준 가입 후 첫 6개월 후에, 그리고 그 이후 매년 COUNTER가 제시한 실행지침을 철저히 준수하는지를 엄격하게 감사하고 이를 통과하였을 때에만 COUNTER를 준수하고 있다는 표시의 'COUNTER compliant' 용어를 사용할 수 있도록 허용하고 있다. 그리고 이 결과를 COUNTER 웹사이트의 'Compliant Vendors' 페이지에 제공하고 있으므로 이 벤더 목록은 매년 변동되는 것이 특징이다. 그런데 COUNTER 표준을 준수하는 것으로 감사받았던 벤더들 중에서 새로운 실행지침이 효력을 발생한 이후에 추가적인 감사 과정 없이 여전히 'COUNTER 준수 벤더'(COUNTER compliant vendor)라고 자신들의 웹사이트에 홍보하고 있는 문제점이 종종 발견되기도 하였

다. Bernd-Christoph Kaemper의 조사에 따르면³⁾ 2005년에 COUNTER 준수 벤더로 감사받았던 14개의 기관 중에서 절반에 해당하는 7개의 기관이 2007년 COUNTER 준수 벤더 목록에 빠져있음에도 불구하고 여전히 자신들의 웹사이트에서 자신들을 COUNTER 표준을 준수하는 벤더라고 홍보하고 있다고 지적한 바 있다. 그러므로 도서관 및 컨소시엄 관계자는 이용통계의 원활한 공급을 위하여 계약하고자 하는 벤더가 COUNTER 표준을 준수하는지에 대한 여부를 정확히 확인해볼 필요가 있다.

또한 COUNTER 3판을 준수하는 벤더의 수가 급증한 데에는 COUNTER 이용통계 보고서를 자동 생성하여 제공해주는 웹 분석 솔루션의 개발이 많은 영향을 미쳤다고 할 수 있다. 현재 Atypen, HighWire Press, ISSEL, Macmillan Publishing Solutions(MPS Insight), MetaPress, Project Euclid, Scholarly IQ의 7개 기관에서 벤더 및 출판사를 대신하여 COUNTER 3판의 실행지침을 준수하는 이용통계 보고서를 제공하는 서비스를 제공하고 있다.

예를 들어, MPS Insight는 Macmillan의 자회사인 MPS Technologies가 벤더 및 출판사를 위해 COUNTER 3판이 요구하는 모든 조건을 충족시키는 SUSHI 기반의 이용통계 보고서를 자동으로 생성하여 제공해주는 신규 서비스이

3) 각주 2)에서 언급된 LIBLICENSE-L의 메일, <<http://www.library.yale.edu/~license/ListArchives/0703/msg00060.html>>.

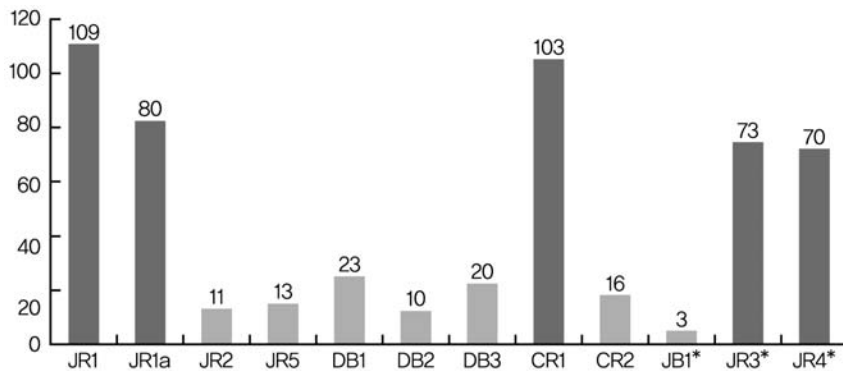
다4). 이 서비스는 이용통계 보고서가 벤더 및 도서관의 로컬 시스템에서 자동으로 검색되도록 하며, 사용하기 쉽고 맞춤화된 인터페이스를 제공하고 있다. Information Today에 실린 기사⁵⁾에 따르면, 이 서비스의 베타테스트에 초대된 사서들 중 96% 이상이 MPS Insight가 이용하기 쉽거나 매우 쉽다고 평가하였다고 한다. 결국 벤더들이 개별적으로 COUNTER의 표준에 응하기보다 이용통계 집계 및 보고서 생성에 필요한 기술을 보유한 별도의 기관에서 이용통계 솔루션을 제공하는 것이 이용통계 활성화에 보다 현실적인 방안이 될 수 있을 것으로 본다.

3.3 이용통계 보고서 제공 현황

COUNTER 3판에서 제시한 이용통계 보고

서의 유형은 추가 옵션 이용통계 보고서인 JB1, JR3, JR5까지 모두 포함하여 총 12개였다. COUNTER 3판을 준수하는 113개 벤더가 현재 도서관 및 각 회원 기관에게 제공하는 이용통계 보고서의 종류를 합하여 본 결과 총 531건이었다. 즉, 각 벤더별로 평균 4.7개 유형의 이용통계 보고서를 제공하고 있었다.

유형별로 제공되는 현황을 살펴보면(〈그림 2〉 참조), 거의 대부분의 벤더가 저널 이용통계 보고서인 JR1(109기관, 96.5%)과 컨소시엄 이용통계 보고서인 CR1(103기관, 91.2%)을 제공하고 있다. 그리고 아카이브 이용통계 보고서인 JR1a(70.8%), 추가 옵션 사항인 저널 보고서 JR3(64.6%)와 JR4(61.9%)가 그 다음으로 많이 제공되었다. 즉, 벤더가 제공하는 평균 4.7개의 이용통계 보고서는 주로



(*는 추가 옵션 보고서를 의미함)

〈그림 2〉 보고서 유형별 출판사 제공 현황

4) <<http://www.mpstechnologies.com/insight.php>>. [인용 2010. 8. 4].

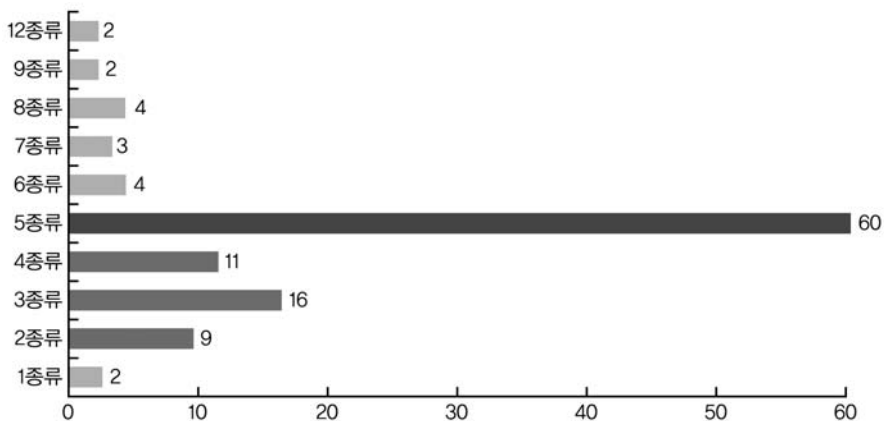
5) "MPS Insight Provides New Usage Reporting Service for Publishers." 2009. *Information Today*, February 26. [인용 2010. 8. 4]. <<http://newsbreaks.infotoday.com/Digest/MPS-Insight-Provides-New-Usage-Reporting-Service-for-Publishers-52791.asp>>.

JR1, JR1a, CR1, JR3, JR4를 의미함을 알 수 있다.

이 중에서 JR3과 JR4는 벤더가 옵션으로 선택할 수 있는 추가적인 보고서임에도 불구하고 현재 타 유형의 보고서에 비해 많이 제공되고 있다는 사실은 벤더가 전자저널 이용통계의 중요성을 인지하고 저널별/월별 원문요청 건수(JR1) 및 저널별/월별 아카이브 원문요청 건수(JR1a) 이외의 상세한 이용통계를 제공하는데 활발히 노력하고 있음을 알려준다. 이러한 추가적인 이용통계 보고서는 도서관 및 회원기관에게 원문요청 건수 이외에 페이지 유형에 따른 요청건수 및 접속거절 건수, 서비스별 검색건수 등의 자세한 데이터를 제공해 줌으로써 도서관 이용자의 전자저널 이용행태를 보다 풍부하게 파악할 수 있게 해줄 뿐만 아니라, 해당 벤더의 전자저널 이용 현황을 파악하는 데 핵심적인 자료로 활용될 수 있기 때문이다.

각 벤더들이 제공하는 이용통계 보고서의 개수를 자세히 살펴보면 <그림 3>과 같다. 5종류의 보고서를 제공하는 벤더가 60곳으로 전체 53.1%를 차지하였으며, 이때에는 앞서 언급하였듯이 JR1, JR1a, CR1, JR3, JR4의 보고서를 제공하는 것이 일반적이었다. 그 다음으로 3종류(14.2%), 4종류(9.7%), 2종류(8.0%)의 보고서가 제공되는 경우가 전체의 30%를 차지하였다. 3종류의 이용통계 보고서를 제공하는 경우에는 JR1, JR1a(또는 JR5)의 저널 보고서 2개와 CR1의 컨소시엄 보고서를 생성하였고, 4종류의 보고서를 제공하는 경우에는 JR1과 CR1 이외에 DB 보고서 중에 하나, 그리고 추가 옵션 보고서인 JR3, JR4 중에 하나를 제공하는 편이었다. 물론 2종류의 보고서를 제공하는 경우에는 JR1과 CR1이 주로 제공되었다.

또한 6종류 이상의 많은 보고서를 제공하는 경우(15기관, 13.3%)에는 Emerald, ProQuest,



<그림 3> 제공되는 이용통계 보고서 개수 현황

〈표 2〉 6개 이상의 이용통계 보고서를 제공하는 벤더 목록

벤더	보고서 개수
ISSEL*	12
Scholarly iQ*	12
American Institute of Physics	9
Wolters Kluwer Health Medical Research	9
Emerald Group Publishing	8
OVID Technologies	8
ProQuest	8
RMIT Publishing	8
Atypon* (excluding Atypon Link)	7
Chadwyck-Healey	7
Macmillan Publishing Solutions (MPS Insight)*	7
American Association for the Advancement of Science	6
CSA	6
HighWire Press*	6
OCLC	6

- 별표(*) 표시는 COUNTER 3판 이용통계 보고서 제공 솔루션 서비스 기관

OCLC 등의 규모가 큰 벤더가 해당되었다(〈표 2〉 참조). 그리고 이 중에서 12개의 모든 보고서를 제공하는 ISSEL과 Scholarly iQ, 7개의 보고서를 제공하는 Atypon, Macmillan Publishing Solutions(MPS Insight), 6개의 보고서를 제공하는 HighWire Press의 5곳은 다른 벤더를 대신하여 SHUSI 기반의 COUNTER 이용통계 보고서를 생성하여 제공해주는 솔루션 서비스를 제공하는 기관이었다.

4. 전자저널 관리를 위한 COUNTER 이용통계 활용 방안

이용데이터에 대한 깊이 있는 분석이 자관

이용자의 전자저널 이용을 잘 설명해줄 수 있을 테지만, 사서들이 현장에서 복잡한 통계치를 활용하여 심도 있는 분석을 실시하는 것이 현실적으로 쉽지 않다는 점을 생각해야 한다. 특정대상 도서관의 일정기간 이용데이터를 심층 분석하는 프로젝트를 진행하여 전자저널 이용에 대한 큰 그림을 그리는 것은 분명 학계 및 현장의 협력으로 진행되어야 할 필요가 있는 향후 과제이다. 그러나 현 단계에서는 이용통계 분석이라는 새로운 업무가 기존의 업무에 추가된 상황에서 많은 욕심은 오히려 금물이 될 수 있다. 또한 벤더가 제공하는 이용통계 보고서를 가능한 많이 변경하지 않은 채 분석하는 것이 가장 시간낭비적인 요소들을 방지할 수 있다. 그러므로 본 연구에서는

COUNTER 표준에 제시되는 기본 데이터 요소를 바탕으로 핵심 통계치를 제안하고, 사서들이 보다 쉽게 분석하고 활용할 수 있는 방안을 제시하고자 한다.

4.1 실무에 유용한 통계치 제안

COUNTER 이용통계 보고서를 통해 제공되는 데이터 요소는 원문요청 건수, 접속거절수, 접속건수, 검색건수 4가지였다. 전자저널의 경우 JR1을 통해 저널별/월별 원문요청 건수가 가장 기본적으로 제공되었고 전자저널 패키지가 검색 데이터베이스의 역할을 함께 하는 경우에는 DB1 또는 DB3를 통해 데이터베이스 접속건수와 검색건수, 그렇지 않은 경우에는 JR4를 통해 서비스별/월별 검색건수, 또는 JR3를 통해 페이지 타입별 요청 건수를 제공하였다. 그러나 JR1의 제공률이 96.5%로 저널별 원문요청 건수가 거의 모든 벤더로부터 제공되는 반면, JR3와 JR4의 제공률은 각각 64.6%, 61.9%이었으며 그 이외의 제공률은 JR2 9.7%, DB1 20.4%, DB3 17.7%로 저조한 편이었다. 즉, COUNTER가 제시한 4개의 데이터 요소 중에서 전자저널의 이용을 분석하는 데 가장 핵심이 되는 값은 여러 벤더의 이용통계 보고서에서 가장 쉽게 통합되어 비교할 수 있는 ‘원문요청 건수’이다. Baker & Read(2008)도 미국 및 캐나다 연구도서관 사서들이 전자저널 이용 데이터를 활용하는데 원문요청 건수가 가장 유용한 값이라고 응

답하였다고 밝힌 바 있다.

이와 달리 Hahn & Faulkner(2002)는 전자저널 이용통계에서 가장 중요한 요소로 ‘논문 수’를 꼽았다. 국내 연구에서도 이를 이용한 전자저널 이용통계 분석이 종종 이루어진 바 있으나, 앞서 살펴보았듯이 COUNTER 이용통계 보고서에서는 논문 수가 제공되지 않고 있다. 결국 논문 수를 분석에 활용하기 위해서는 저널별 또는 패키지별로 포함된 논문의 양을 수작업으로 파악해야 하는 현실적인 어려움이 있다. 이론적으로는 논문 수를 고려하여 원문요청 건수를 파악하는 것이 보다 정교할 수 있지만 현실적으로는 실무에 적용하는 데 부적합한 것이 사실이다. 따라서 본 장에서는 JR1의 원문요청 건수를 이용통계 분석의 핵심값으로 사용하고자 한다.

이용통계 분석의 수준은 도서관에서 구독 중인 패키지 전체에 대한 분석과 개별 패키지 내 저널별 이용에 대한 분석으로 구분될 수 있는데, 본 연구에서는 저널별 통계 분석보다는 패키지별 통계 분석에 보다 초점을 맞출 것이다. 과거의 인쇄환경이라면 저널별로 그 이용값을 비교하는 것이 당연했을 테지만 지금의 전자저널은 출판사에서 생산하는 수십 또는 수백 종의 저널을 하나로 묶어 패키지로 계약하고 패키지별로 한 개의 보고서를 제공받는 것이 일반적이다. 만약 여러 벤더로부터 전달받은 플랫폼별, 출판사별, 또는 서비스별 이용통계 보고서를 가지고 개별 저널을 분석하고자 한다면 사서는 수천 종에 이르는 저널 각각의 이용 데

이터를 별도로 통합해야 하는 작업에 긴 시간을 할애해야만 한다. 이를 보다 간단히 하기 위해 MS Access 등의 관계형 DB 프로그램을 활용할 수도 있다. 하지만 많은 시간과 노력을 들여 저널별 이용통계를 분석하였다고 하더라도 전자저널의 라이선스 계약은 여전히 패키지별로 이루어지고 있기 때문에 효율적인 전자저널 관리를 위해서는 저널별이 아닌 패키지별 분석이 보다 합리적이라는 결론이 도출된다.

이러한 내용을 전제로 본 연구에서 전자저널 이용통계 분석을 위해 제안하는 통계치는 크게 저널별 원문이용도, 저널이용률, 원문이용단가, 고이용군(most use), 저이용군(low use)이다. 분석의 수준에 따라 투입되는 변수가 달라질 수는 있겠으나 각 통계치는 <표 3>과 같이 구해질 수 있다. 이 때, 고이용 및 저이용 저널군을 구하기 위한 기준은 자관의 구독 패키지 수 및 전자저널 수를 고려하여 설정될 수 있다. 한 예로, 전자저널의 경우에도 인쇄저널과 마찬가지로 20:80의 로트카 법칙이

적용되거나(Conyers & Dalton 2005), 47%에 해당하는 저널이 전체 이용의 80%를 차지하였다(Evans & Peters 2005)는 연구결과에 근거하여 20~40%의 범위 내에서 고이용군으로 설정할 수 있을 것이다.

4.2 통계치의 효과적 활용 방안

4.2.1 활용 준비작업

COUNTER를 준수하는 벤더들은 앞에서 살펴본 <그림 1>과 같은 형태로 각 플랫폼에서 제공하는 저널별/월별 원문요청 건수를 JR1을 통해 제공한다. 컨소시엄을 통해 전자저널을 공동구매하는 경우에도 개별 기관별 원문요청 건수를 CR1을 통해 제공하고 있다. 따라서 전자저널 이용통계 관리를 담당하는 사서는 도서관 전체의 그림을 그리기 위해서 새로운 스프레드시트에 도서관에서 구독하는 패키지 전체를 한 눈에 분석할 수 있도록 패키지(DB 또는 플랫폼)별 이용통계를 요약 정리하는 준비 작업이 필요하다.

본 연구에서는 벤더들의 제공률이 가장 높은 원문요청 건수를 전자저널 이용통계의 핵심 데이터로 보기 때문에 패키지별 원문요청 건수를 기본으로 사용하며, 더 나아가 패키지별 원문요청 건수에 대한 깊이 있는 분석을 위해 패키지에 포함된 구독 저널 수, 이용 저널 수, 구독 가격의 3가지 변인을 추가하였다. 이외에 DB보고서 및 옵션 저널보고서에서 제공

<표 3> 전자저널 이용 분석을 위한 핵심 통계치

<ul style="list-style-type: none"> • 저널별 원문이용도 = $\frac{\text{총 원문요청 건수}}{\text{구독 저널 수}}$ • 저널이용률(%) = $\frac{\text{1번 이상 이용된 저널 수}}{\text{구독 저널 수}} \times 100$ • 원문이용단가 = $\frac{\text{구독비용}}{\text{원문요청 건수}}$ • 고이용군 = 원문요청 건수 상위 ◇ ◇ % • 저이용군 = 원문요청 건수 하위 ◇ ◇ %

하는 검색건수, 접속건수가 추가 변인으로 요약되어 분석되어 질 수 있으나 본 연구에서는 생략하기로 한다.

보다 쉬운 이해를 위해 전자저널 패키지 5개에 대한 요약 스프레드시트를 <그림 4>와 같이 가상으로 제시하였다. B열의 원문요청 건수는 JR1에서 전체저널에 대한 해당연도 합계(YTD total)에 해당한다(예: <그림 1>의 I6). C열의 ‘구독 저널 수’는 각각의 JR1에 수록된 저널의 종수를 구하면 된다. JR1에서 개별 저널명은 7행에서부터 시작되므로 저널명의 마지막 행번호를 활용하여 전체 저널 수를 계산할 수 있다⁶⁾. D열의 ‘이용 저널 수’는 구독하는 저널 중 최소 1번 이상 이용된 저널 수를 기록한다. 이 때 해당연도의 원문요청 건수 합계가 0인 저널을 제외한 종수를 구하기 위해 지정한 범위 내에서 조건에 맞는 셀의 개수를 구해주는 COUNTIF() 함수를 활용하는 것이 좋다. 예를 들어 JR1에 저널별 원문요청 건수 합계가 R7에서 R112까지 존재하고 이 범위 내에서 1 이상의 값을 갖는 셀의 개수를 구

하기 위해서는 ‘=COUNTIF(R7:R112, “>=1”)’이라고 입력한다. E열의 구독가격은 도서관에서 구독하는 전자저널 패키지의 가격을 입력하면 된다.

주어진 조건대로 모든 값을 입력하고 나면 가장 마지막 행에는 <그림 4>의 7행처럼 각 열의 평균값을 AVERAGE() 함수를 이용하여 구한다. 평균은 값의 집중도를 살펴보는 기술 통계의 기본 통계치로, 각 패키지의 값을 분석하는 데 기준이 될 것이다. 학교에서 성적을 부여할 때 학생들의 성적을 타 대학 또는 타 학과 학생들과 비교하는 것이 아니라 같은 반 학생의 점수만을 비교하여 평균, 1등, 순위 등을 구한다. 이처럼 전자저널 이용통계에서도 자관 도서관에서 구독 중인 패키지별 원문이용 현황을 통해 성적을 부여한다고 이해하면 쉬울 것이다. 이를 통해 기본적으로 자관 도서관의 평균 원문요청 건수는 어느 정도이고, 평균보다 원문이 많이 이용된 패키지는 무엇이며, 평균보다 적게 이용된 패키지는 무엇인지를 파악할 수 있다.

	A	B	C	D	E
1	전자저널 패키지	원문요청 건수	저널 수	이용 저널 수	구독가격(만원)
2	A	50,000	50	45	5,000
3	B	3,000	200	100	1,500
4	C	10,000	400	300	7,500
5	D	50,000	1,000	950	1,500
6	E	30,000	20	20	1,500
7	평균	28,600	334	283	3,400

<그림 4> 전자저널 패키지별 이용통계 분석을 위한 스프레드시트 작성 예시

6) 범위 내에서 비어있지 않은 셀의 개수를 구하는 COUNTA() 함수를 이용할 수도 있다. 예를 들어, 저널이 기록된 마지막 행번호가 112였을 때 =COUNTA(A7:A112)라고 입력하면 106개의 저널명이 있는 것으로 계산된다.

4.2.2 패키지의 이용통계 분석

① 저널별 원문이용도

저널별 원문이용도는 각 패키지에 포함된 구독 저널 종수를 고려하여 원문요청 건수를 해석하고자 하는 것으로 JR1에서 제공한 원문요청 건수(예: B2)를 구독 저널 수(예: C2)로 나눈 값이다. 따라서 <그림 5>의 F2에 ' $B2/C2$ '라고 계산식을 입력하면 한 개의 저널당 평균적으로 몇 건의 원문요청이 있었는지를 파악할 수 있다. 예제에서는 전체적으로 한 개의 저널에서 평균 86건의 원문 다운로드가 이루어졌고, 패키지 E와 패키지 A의 저널별 원문이용도가 타 패키지와 비교할 수 없을 정도로 높았으며 그 외의 패키지는 평균보다 적게 이용되고 있었다. 특히 패키지 B에 수록된 전자저널이 가장 저조하게 이용되는 경향을 보인 것을 파악할 수 있다.

패키지의 저널별 원문이용도가 높다는 것은 해당 패키지가 그만큼 자관 이용자의 정보요구를 잘 충족시키고 있다고 볼 수 있다. 반면에 낮은 저널별 원문이용도는 자관 이용자의 정보요구와 적합하지 않은 전자저널로 구성되어 있음을 암시한다. 즉, 저널별 원문이용도는

도서관에서 구독 중인 패키지의 전자저널 구성이 얼마나 효과적인지를 분석하는 지표로 활용될 수 있다.

② 저널이용률

저널이용률은 구독한 전체 저널 중에서 한번 이상의 원문 이용이 있는 전자저널의 분포를 구하는 통계치이다. 따라서 구독 저널 수(예: C2)와 이용 저널 수(예: D2)를 이용하여 <그림 5>의 G2에는 ' $D2/C2*100$ '이라고 입력한다. 패키지에서 제공하는 모든 저널이 한번 이상 이용되었을 때 저널이용률은 100%가 나오며 모든 저널이 한번도 이용되지 않았을 때에는 0%의 값을 갖게 된다. 예제에서는 패키지 E의 저널 구성이 가장 효과적인 것으로 나타났으며, 패키지 D와 패키지 A의 경우 평균 저널이용률보다 높게 나타남으로써 구독하는 저널이 효과적으로 이용되고 있음을 알 수 있다. 반면에 패키지 B의 경우 저널이용률이 50%로 구독저널의 절반 이상이 전혀 이용되지 않고 있어 가장 효과적이지 못한 것으로 분석되었다.

이처럼 저널이용률의 전체 평균이 100%에

	A	B	C	D	F	G
1	전자저널 패키지	원문요청 건수	저널 수	이용 저널 수	원문이용도	저널이용률(%)
2	A	50,000	50	45	1,000	90
3	B	3,000	200	100	15	50
4	C	10,000	400	300	25	75
5	D	50,000	1,000	950	50	95
6	E	30,000	20	20	1,500	100
7	평균	28,600	334	283	86	85

<그림 5> 저널별 원문이용도, 저널이용률을 이용한 분석 예

가까울수록, 그리고 패키지별 저널이용률이 100%에 가까울수록 현재 구독 중인 전자저널의 구성이 이용자의 요구를 잘 반영하고 있으며 비효율적인 예산 낭비가 적다고 해석될 수 있다. 그러나 현재 이용이 없다하여 이용자의 잠재적 요구가 없다고 단정 지을 수는 없다. 또한 지금의 전자저널 구독 환경은 일반적으로 이용이 존재하는 저널만을 선별하여 구독할 수는 없는 상황이다. 그러므로 Bleicic, Fiscella, & Wiberley(2001)는 도서관이 자관의 실정에 맞게 “몇 %의 이용이 이루어지길 원하는가?”와 같은 기준을 정해 구독 및 취소 정책을 마련해야 한다고 조언하고 있다. 예를 들어 저널이용률이 50% 이하일 때 전자저널의 패키지 구성에 대해 재협상을 시도한다 등의 내부 기준이 마련될 필요가 있다.

③ 원문이용단가

원문이용단가는 이용자의 원문요청 건수를 고려하여 전자저널 패키지 구독가격을 해석하기 위한 값으로, 구독가격(예: <그림 6>의 E2)을 원문요청 건수(예: B2)로 나누어주면 된다. 예제 속의 도서관은 한 건당 평균 1,189원을

지불하여 전자저널 원문을 다운로드 하고 있는 것으로 나타났다. 평균을 기준으로 고려할 때, 패키지 D와 패키지 E에서 제공하는 전자저널은 한 건당 300원, 500원으로 매우 저렴한 편이었으나 패키지 B와 C는 한 건당 5,000원, 7,500원으로 평균보다 높은 가격에 제공되고 있음을 알 수 있다.

원문이용단가가 높게 나온 경우에는 패키지의 특성을 기반으로 그 원인을 분석해볼 필요가 있다. 높은 단가를 설명할 이유가 부족할 경우 이용통계 분석결과를 라이선스 갱신 시 가격 재협상 및 구독 갱신에 대한 의사결정 근거자료로 활용할 수 있다. 간혹 타 도서관의 원문이용단가와 비교해야 할 것 같다는 생각이 들기도 하겠지만, 라이선스 구독은 엄연히 협상이라는 단계를 거쳐 이루어지는 것이기 때문에 표면적으로 공개된 가격에는 큰 차이가 없을 것으로 판단된다. 다만 자관의 이용이 적다는 이유로 무조건 저널의 가격 인하를 요구할 수는 없기 때문에 구독을 지속할 것인지 혹은 국내 컨소시엄을 통한 가격협상이 가능한지를 고려해봐야 할 필요가 있다. Conyers & Dalton(2005)의 조사에 따르면 높은 가격대의 저널들이 많은 이용

	A	B	C	E	F	H
1	전자저널 패키지	원문요청 건수	저널 수	구독가격(만원)	원문이용도	원문이용단가(원)
2	A	50,000	50	5,000	1,000	1,000
3	B	3,000	200	1,500	15	5,000
4	C	10,000	400	7,500	25	7,500
5	D	50,000	1,000	1,500	50	300
6	E	30,000	20	1,500	1,500	500
7	평균	28,600	334	3,400	86	1,189

<그림 6> 원문이용단가를 이용한 분석 예

을 보이고 낮은 가격대의 저널들은 상대적으로 저조한 이용을 보임으로써 모든 도서관에서 유사한 이용대비 단가를 보이는 것으로 분석되었다. 따라서 타 도서관과의 비교를 목적으로 하 기보다 자관에서 수용 가능한 원문이용단가가 어느 수준인지를 우선적으로 결정할 필요가 있다(Kraemer 2006).

④ 고이용군/저이용군

고이용군과 저이용군을 추출하기 위해 <그림 7>과 같이 엑셀의 자동필터 기능을 이용 하면 간단하다. 저널별 원문이용도를 기준으로 상위 및 하위 몇 %에 해당하거나 원하는 개수(항목)에 해당하는 값을 검색하여 그 결과 값을 제공하기 때문이다. 이 때에 B열의 단순한 원문요청 건수보다 패키지에 포함된 구독 저널 수를 고려하여 계산된 F열의 저널별 원문이용도를 활용하는 것이 바람직하다. 예를 들어, <그림 7>에서는 패키지 A와 E가 상위 40%의 높은 이용도를 보이는 패키지로 검색

되었다. 그러나 만약 원문요청 건수로 추출한다면, 5만 건의 이용을 보인 패키지 A와 D가 검색될 것이다(<그림 6> 참조). 그러나 패키지 E의 경우 전체 원문요청 건수는 3만 건으로 패키지 D보다 작지만 저널별 원문이용도가 50인 패키지 D에 비해 활발히 이용된 것을 알 수 있다.

본 연구에서는 지면 제약에 따른 이유로 패키지별 분석 방법만을 보여주고 있으나 이 방법은 각 패키지내의 전자저널별 분석에 쉽게 적용할 수 있다. JR1의 각 저널별 원문요청 건수의 당해연도 합계 항목을 이용하여 상위 20~40%에 해당하는 저널군을 식별할 수 있으며 각 패키지별로 분석된 고이용 저널군, 저이용 저널군에 대한 전자저널 관리 방안을 마련할 수 있다.

많이 이용되는 패키지 또는 저널의 경우는 이용자의 접근이 많은 만큼 보다 쉽게 해당 자료로 이동할 수 있도록 인터페이스를 제공함으로써 이용자의 편의를 제공하는 방법을 고



<그림 7> 자동필터 기능을 이용한 상위 40% 검색 예

7) 엑셀(2007 버전 기준)의 메뉴 [데이터]→[필터]→필터의 기준이 될 제목열(저널별 원문이용도)의 역삼각형 단추 클릭→[숫자필터]→[상위 10]→상위 또는 하위의 몇 % 또는 몇 개의 항목을 검색할 것인지 검색조건을 입력한다.

려할 수 있다. 해당 패키지 및 해당 저널로의 접근을 용이하도록 하기 위해 도서관 웹페이지 첫 화면에 '인기저널 목록'을 공지하거나 '배너'를 삽입하여 단축경로를 마련하는 것도 한 방법이다. 전자저널의 경우에도 인쇄저널과 마찬가지로 상대적으로 적은 중수가 높은 이용을 차지하고 있으므로 상위 20~40%에 해당하는 핵심 저널 및 패키지를 파악하고 이들에 대한 이용을 보다 자극하여 전자저널의 이용을 활성화시킬 필요가 있겠다.

이용이 전혀 없거나 이용이 저조한 경우는 그 원인을 분석하는 것이 우선적으로 필요하겠으며 동시에 전자저널 구독조건을 다른 방식으로 협상하거나 취소하는 방안을 고려할 필요가 있다. 또한 반대로 저이용군의 이용률을 높이기 위해 역시 이용자의 눈에 쉽게 띄 수 있는 접근 경로를 제공하거나 해당 DB 및 저널을 대상으로 이용자 교육을 확대할 필요가 있다. 만약 저널이 새로 발간된 것이거나 혹은 도서관이 처음으로 구독하기 시작한 저널의 경우에는 아무래도 이용자에게 친숙하게 되기까지 시간이 필요하다. Bordeaux, Kraemer, & Sullenger(2005)는 적어도 이용자가 새로운 전자저널에 처음으로 노출된 후 이들을 지속적으로 이용하기까지 3개월에서 6개월이라는 물리적인 준비기간이 필요하기 때문에 도서관은 이러한 새로운 저널들에 대한 홍보 노력을 지속해야 한다고 하였다.

5. 결론 및 제언

본 연구는 전자저널 이용통계와 관련된 국내외 선행연구 분석과 COUNTER 3판의 전자저널 이용통계 제공 현황 분석을 기반으로 전자저널 관리의 효과를 극대화시킬 수 있는 이용통계 분석 및 활용 방안을 모색하고자 하였다. 특히 본 연구는 국내 도서관 현장에서 쉽게 이용통계를 수집하고 분석하는 데 참고할 수 있는 가이드라인을 제시하는 데 초점을 두었다.

본 논문에서 제시된 이용통계의 활용은 도서관 이용자의 전자저널 이용행태를 파악하는데 도움을 줄 것이며 전자저널 관리와 관련된 도서관 운영의 의사결정 자료로 활용 가능할 것이다. 도서관을 경영하는 데 있어서 제한된 자원의 효율적인 배분을 위해 정확하고 신뢰할만한 통계 및 자료가 필수적이다. 전자저널 이용통계 데이터를 기반으로 장서관리의 효과성과 비용의 효율성 분석이 가능하며 고이용/저이용 저널군 등을 비롯하여 다양한 자관의 전자저널 이용 패턴을 파악함으로써 이에 맞춘 향상된 서비스를 제공할 수 있을 것으로 기대된다.

다만 본 논문은 사서 실무자들의 효과적인 활용을 위해 산출에 어려움이 있는 저널당 논문 수를 범위에서 제외시켰으므로 향후 연구에서는 이를 고려할 필요가 있겠으며 제시한 이용통계치 활용에 대한 검증 및 이용통계의 또 다른 지표인 영향력 지수와 그 변종 등을

추가로 검토할 필요가 있겠다.

끝으로 전자저널 이용통계의 활성화를 위해 보다 논의되고 고려되어야 하는 몇 가지 이슈들을 제언하고자 한다. 첫째, 도서관은 어떤 통계와 어떤 포맷이 자관의 요구에 최적한지 정의하고 이를 어떻게 활용할 것인지에 대한 전략을 마련할 필요가 있다. COUNTER 표준안에 대한 이해가 바탕이 된 후, 자관의 목적에 부합시킬 통계치가 무엇이며 이에 대한 활용 방안을 모색해야 한다. 자관의 특성을 고려하여 이용통계 분석의 기준을 정립하고 이를 전자저널 구독 및 관리의 의사결정 근거자료로 이용해야 한다. 이를 위해 이용통계 활용 사례를 공유하고 이용통계 분석 방법을 교육하는 포럼 및 워크숍 등의 프로젝트가 진행될 필요가 있다. 영국의 경우에는 사서들의 이용통계 분석을 지원하기 위해 참가비를 내고 이용통계 분석에 관한 전반적인 노하우를 교육 받고 서로의 경험을 공유하는 프로젝트를 운영한 사례가 있다(Conyers & Dalton 2005). 이에 반해 국내에서는 아직 이용통계 활용에 대한 적극적인 활동이 본격적으로 진행되지 않고 있다.

둘째, 이용통계 분석은 월 단위로 진행함으로써 분석된 자관 이용자의 전자저널 이용 행태를 도서관 홍보, 이용자 교육, 장서관리 등의 전 영역에 즉각적으로 반영할 필요가 있다. 분명 매월 여러 벤더로부터 제공되는 데이터를 수집하고 통합 및 정제하는 것은 상당한 시간을 요구한다. 빙햄턴 대학도서관 사서인 Bo-

rdeaux는 웹에서 데이터를 가져와 DB용, 저널 패키지용 스프레드시트를 하나씩 만들며, 검색건수, 원문요청 건수를 계산하고 예외적인 사항을 기록하는 데 일주일에 5시간, 한 달에 2주, 즉 월평균 10시간을 이용통계 다운로드 및 입력에 소비한다고 하였다(Bordeaux, Kraemer & Sullenger 2005). 그러나 전자저널 이용 현황을 장서관리 및 도서관 운영 정책에 즉각적으로 반영시키기 위해서는 보다 적극적이고 꾸준한 노력이 필요하다. 이를 위해 분석된 이용통계는 도서관 인트라넷에 공개되어 전 직원이 공유할 필요가 있다. 또한 이용통계 전담사서가 이용통계 수집 및 분석에 충분한 시간을 할애할 수 있도록 업무 조정이 이루어질 필요도 있겠다.

셋째, 이용통계 분석이 어느 정도 안정화되면 이용통계 전담사서는 분석의 깊이를 보다 확대할 필요가 있다. 예를 들어 접속건수 및 검색건수를 통해 DB 및 전자저널의 이용 행태를 살펴볼 수 있으며 주제분야에 따른 저널별 원문이용도의 차이를 분석할 수 있다. 또한 지속적으로 이용데이터를 수집하고 통계 분석 결과를 저장하여 향후 증감률을 모니터링하고 분석하는 작업이 이루어져야 한다. 자관의 전자저널 이용에 대한 월별, 분기별, 연도별 변화를 살펴봄으로써 그 원인을 파악하고 이를 장서관리 정책에 반영시킬 필요가 있다.

넷째, COUNTER 보고서의 이용통계 산출 기간이 회계연도, 학년도, 또는 원하는 기간 등과 같이 유연하게 조정하여 검색할 수 있도록

록 변화가 요구된다. 현재 COUNTER 보고서는 회계연도를 기준으로 1월부터 12월까지의 월별 이용통계와 해당년도 합계를 제공하고 있다. 그러므로 3월부터 다음해 2월까지의 학년도로 운영되는 국내 대학도서관의 경우에는 2년간의 이용통계를 바탕으로 재집계해야 하는 번거로움이 있다. 이용통계 집계, 정제, 분석 과정을 단순화하기 위해 보고서 검색 및 제공 방식의 유연성이 보다 확보되어야 할 것이다.

참고문헌

- 김혜선. 2004. 전자저널 이용통계서비스의 품질 평가에 대한 연구. 『정보관리연구』, 35(4): 35-56.
- 박은경. 2005. 『대학도서관의 패키지형 전자저널 수서정책 개선방안에 관한 연구』. 석사학위 논문, 이화여자대학교 대학원, 문헌정보학과.
- 심원식. 2005. 전자정보 이용통계 활용 전략. 『정보관리학회지』, 22(2): 5-21.
- 정은경. 2008. 『전자저널 이용통계 표준화에 관한 연구』. 석사학위 논문, 이화여자대학교 대학원, 문헌정보학과.
- 황옥경. 2007. 대학도서관에서의 전자저널 이용통계 제공 및 활용 현황. 『정보관리연구』, 38(4): 68-87.
- Baker, G. and E. J. Read. 2008. "Vendor-supplied usage data for electronic resources: a resources: a survey of academic libraries." *Learned Publishing*, 21(1): 48-57.
- Blecic, D. D., J. B. Fiscella, and S. E. Wiberley, Jr. 2001. "The Measurement of Use of Web-based Information Resources: An Early Look at Vendor-supplied Data." *College & Research Libraries*, 62(5): 434-453.
- Blecic, D. D., J. B. Fiscella, and S. E. Wiberley, Jr. 2007. "Measurement of use of electronic resources: advances in use statistics and innovations in resource functionality." *College & Research Libraries*, 68(1): 26-44. [cited 2010.01.19].
 <<http://cr.l.acrl.org/content/68/1/26.full.pdf+html>>.
- Bordeaux, A., A. B. Kraemer, and P. Sul-lenger. 2005. "Making the Most of Your Usage Statistic." *The Serial Librarian*, 48(3): 295-299.
- Conyers, A. and P. Dalton, 2005. *NESLI2 Analysis of Usage Statistics, Summary Report*. [cited 2010.01.27].
 <<http://www.ebase.bcu.ac.uk/docs/jiscnesli2summaryeb.pdf>>.
- COUNTER. 2008a. *Appendix H: Optional Additional Usage Reports*. [cited 2010.01.29].
 <www.projectcounter.org/r3/r3_H>.

- doc).
- _____. 2008b. *Introduction to Release 3 of the COUNTER Code of Practice for Journals and Databases*. [cited 2010.01.29].
 <http://www.projectcounter.org/r3/r3_intro.pdf>.
- _____. 2008c. *The COUNTER Code of Practice – Journals and Databases: Release 3*. [cited 2010.01.29].
 <<http://www.projectcounter.org/r3/Release3D9.pdf>>.
- Evans, P. and J. Peters. 2005. "Analysis of the Dispersal of Use for Journals in Emerald Management Xtra (EMX)." *Interlending & Document Supply*, 33(3): 155-157.
- Hahn, K. L. and Lila A. Faulkner. 2002. "Evaluative Usage-based Metrics for the Selection of E-journals." *College & Research Libraries*, 63(3): 215-227.
- Kraemer, A. 2006. "Ensuring Consistent Usage Statistics, Part 2." *The Serials Librarian*, 50(1): 163-172.
- Luther, J. 2001. *White Paper on Electronic Journal Usage Statistic*. 2nd ed. Washington, DC. Council on Library and Information Resources. [cited 2010.01.19].
 <www.clir.org/pubs/reports/pub94/pub94.pdf>.
- Pesch, O. 2006. "Ensuring Consistent Usage Statistics, Part 1." *The Serials Librarian*, 50(1): 147-161.
- Shim, W. and C. R. McClure. 2002. "Improving database vendors' usage statistics reporting through collaboration between libraries and vendors." *College & Research Libraries*, 63(6): 499-514. [cited 2010.01.19].
 <<http://www.ala.org/ala/mgrps/divs/acrl/publications/crljournal/2002/nov/shim.pdf>>.