

## 학술지 영향력지수와 전자 학술지 이용의 상관관계 분석

### Analysis of the Relationship between Journal Impact Factor and Scholarly e-Journal Usage

김환민, 김재훈, 정도현  
한국과학기술정보연구원

Hwan-min Kim, Jay-hoon Kim, Do-heon Jeong  
Korea Institute of Science & Technology Information

#### 요약

국내에서 해외 학술 전자저널 이용과 학술지 영향력 지수의 관계를 알아보기 위하여 이 연구에서는 컨소시엄 기반으로 사이트 라이선스 형식으로 유통중인 American Chemical Society, Science, Elsevier, Wiley 등 4개 출판사 전자학술지의 2006년 이용량과 ISI의 JCR 2006에 등재된 영향력지수간의 상관관계를 중심으로 분석하였다. 이용량 분석 대상은 국내에서 상대적으로 이용량이 많으면서 이들 4개 출판사의 전자저널을 모두 라이선싱하고 있는 4개 대학과 5개 연구소를 택하였으며, JCR 영향력지수 부여 여부와 출판사, 주제 및 기관유형별로 학술지 영향력지수와 전자학술지 이용의 상관관계를 분석하였다.

#### Abstract

In order to find out the relationship between the indicator of scholarly impact and e-Journal usage in Korea, we analyzed the correlation between the usage statistics of e-Journals published by American Chemical Society, Science, Elsevier and Wiley which was licensed by KESLI consortia in 2006 and impact factor in JCR 2006. We selected 4 universities and 5 research institutes as target institutions which participated in all of the 4 consortia and had higher frequency of e-Journal usage than others. In this correlation analysis we used variables such as the subjects of journals, the types of institution, and impact factor.

## I. 서론

과학, 공학, 의학, 법률, 사회과학, 인문과학 등 영역에서 학술지 이용에 대한 연구는 다수 존재하며 다양한 이유로 수행되어 왔다. 이는 학술지가 학술 연구에 대한 이용자의 요구와 기대를 가장 잘 충족시켜왔기 때문이다. 그러나 학술지 구독료가 전체 정보자료 구입예산에서 차지하는 비중이 매우 높고, 학술지 가격은 도서관 예산 증가율이나 물가상승률과 비교하여 큰 폭으로 증가하고 있기 때문에 학술지 구독을 늘리고자 하는 도서관에 있어 큰 부담으로 작용하고 있다. 따라서 많은 비용을 투입하여 구독하는 학술지가 실제로 잘 활용되고 있는지를 확인하고 수서정책에 반영하는 것은 도서관에 있어 매우 중요하고도 어려운 일이다.

한편 전자출판이 활성화되면서 KESLI 및 KERIS 등 컨소시엄을 통한 전자정보 구독은 대학, 연구소, 대기업, 의료기관 등 학술 연구기능을 수행하는 대부분의 기관으로 이미 확산된 상태다. 또한 국내 연구자들의 전자저널 이용이 꾸준히 증가하고 있으며, 이미 상업 출판사의 전자저널 다운로드 총량은 수 천만건에 달할 정도로 활성화되었다. 그러나 전자학술지 이용 분석은 실제 이용 데이터가 아닌 이용자 설문자료에 기반을 둔 경우가 많다[1]. 전자학술지 이용에 대해 분석할 수 있는 이

용통계 데이터는 출판사에서 각 개별기관에 제공하는 데이터에 의존할 수밖에 없었으므로, 각 기관은 타 기관과 이용정도를 비교할 수 없었을 뿐만 아니라 각 기관의 이용통계가 갖는 의미를 분석하기 어려웠다.

이 연구에서는 KESLI에서 수집한 참가기관의 컨소시엄별 전자학술지 이용통계를 분석하여 impact factor와 전자저널 이용량의 상관관계를 분석하였다. 분석 대상은 ACS (American Chemical Society), ScienceDirect, Science, Wiley 등 4개 컨소시엄이며, 이 컨소시엄에 공통적으로 참가하고 있는 4개 대학과 5개 연구소의 2006년 학술지 타이틀별 이용량에 대한 것이다.

## II. 관련연구

Johan 등은 Tomson의 ISI impact factor를 흉내내어 usage impact factor를 정의하였는데 usage impact factor 순위는 2003년부터 2005년까지 California State University의 이용 데이터를 기초로 산출하였다. 그리고 저널 순위를 일관적인 영향력 지수인 ISI impact factor와 비교하였다. 그 결과 California State University의 이용량에서 도출된 영향력

지수는 해당 커뮤니티의 학문 특성과 인구통계학적 특성을 반영하는 것으로 나타났다[2].

Deborah 등은 웹기반 정보자원 이용 측정 연구를 통해 전자형태의 아티클이 이용자의 컴퓨터에 다운로드 되었다고 해도 그것이 읽히고 면밀히 검토되었다고 할 수 없으며, 컴퓨터 모니터를 통한 이용은 읽는다기보다 브라우징에 가깝다고 주장했다[3]. 하지만 특정 저널이나 아티클의 누적 이용량은 이용자의 관심을 반영하고 있음은 분명하다.

King 등은 과학자들이 특정 아티클과 전체 저널을 읽는 양을 측정하였는데, 전자저널 히트와 다운로드, 트랜잭션 로그, 저널이나 아티클에 인용된 수, 그리고 출판된 아티클의 수와 비교하여 총량을 추정하였다[4].

영국의 교육 및 연구 지원기관인 Joint Information Systems Committee(JISC)는 NESLi2 컨소시엄에 참여하는 출판사로부터 제공된 2004년 1월부터 6월까지의 COUNTER (Counting Online Usage of NeTworked Electronic Resources) 기반 통계자료를 이용하여 17개 기관의 이용수준 및 행태와 경제성을 분석하였는데, 출판사간에 성행하고 있는 '빅딜'에 의해 전혀 이용되지 않는 저널까지 포함한 분석을 수행하였다[5].

이 외에도 국내외의 많은 기관이 전자정보 수집에 참고하기 위하여 이용통계 분석을 통한 경제성 분석이나 이용 효율성 분석 및 이용자의 이용특성 분석을 수행한 연구는 다수 존재하지만 개별기관의 특정 컨소시엄을 대상으로 한 것이 대부분이다.

### III. 자료 수집 및 분석

#### 1. 자료수집 및 처리

##### 1.1 분석 대상 전자학술지

분석 대상인 ACS, Science, ScienceDirect, Wiley 등 4개 컨소시엄은 참가하고 있는 기관이 많아 관종별 비교분석이 가능하고, ISI impact factor 부여 비율이 높아 이용량과의 상관관계 분석 결과에 대하여 타 컨소시엄 보다 의미를 부여할 수 있기 때문이다. 이들 4개 컨소시엄의 2006년 통계데이터에는 ACS 36 종, Science 1 종, ScienceDirect 2,056 종, Wiley 860 종 등 2,953종에 대한 이용통계를 수록하고 있으나 Science Direct와 Wiley간의 타이틀 중복 및 아카이브 타이틀과 커런트 타이틀간의 중복을 제거하고 2,658 종을 추출하였다. 그리고 주제별 분석을 위하여 NDSL 수록 정보와 e-ISSN, p-ISSN 및 학술지 타이틀을 매핑하여 DDC 분류를 추출하였는데, 매핑 되지 않는 ScienceDirect 5 종, Wiley 11 종을 분석 대상에서 제외하고, 2,642종을 분석대상으로 선정하

였다. 표 1에 분석대상 학술지를 컨소시엄별로 나타내었다.

[표 1] 컨소시엄별 분석대상 학술지

구분	ACS	Science	ScienceDirect	Wiley	계
분석대상 종수	36	1	2,038	567	2,642
JCR IF 부여 종수	32	1	1,148	281	1,462
JCR IF 부여 비율	89%	100%	56%	49%	55%
평균 IF	4.98	30.03	2.37	1.98	2.37

##### 1.2 JCR impact factor 추출

분석 대상 학술지 2,642종과 JCR 2006을 대조하여 분석대상 학술지별 impact factor를 추출한 결과 분석대상 학술지의 55%인 1,462종에 대해 대응시킬 수 있었다. 표 1에 보인바와 같이 컨소시엄별로 impact factor가 부여된 학술지는 ACS 32 종(89%), Science 1종(100%), ScienceDirect 1,148종(56%), Wiley 281종(49%) 이었다.

##### 1.3 분석 대상 학술지별 주제분류

분석 대상 학술지 정보와 NDSL 서지정보를 ISSN과 타이틀 매핑에 의해 2,652종의 각 학술지별로 DDC 분류기호를 추출하였으며, 매핑테이블에 의하여 KISTI 표준분류의 대분류로 변환하였다. KISTI 표준분류의 대분류를 다시 그룹핑하여 11개의 주제분류 항목으로 축약하였으며 그 결과를 표 2에 나타내었다.

[표 2] KISTI 과학기술 표준분류의 그룹핑

구분	KISTI 표준분류	대상 종수	IF 종수
건설/건축/환경	AA 건설공학, AB 건축공학, AC 토목공학, AD도시공학, AE 환경공학	84	46
생명과학	BA 생물학, BB 생물공학, BC 생화학, BD 약학, BM 의학	1,176	706
식품/농축수산	BE 수산학, BF 식품, BG 농화학, BH 농림업, BL 축산	79	58
화학/섬유/화공	CA 화학, CB 화공, OC 고분자, CD 요업, CE 섬유	245	171
수학/물리	LA 수학, PA 물리학, PB 진동학	200	142
전기/전자	EC 전기공학, ET 전자공학	49	34
전산/정보	NA 전산학, NB 정보통신, NC 정보학, ND 정보기공, EJ 정보과학	123	67
기계	MA 기계공학, MB 열유체공학, MC 기계제작 및 산업기계, MD 수송공학	105	70
금속/자원/에너지	RA 금속공학, RB 자원공학, RC 에너지공학	129	94
과학기술 일반	SA 과학기술일반	39	18
경제/경영/인문/기타	SB 경영경제, SC 사회과학, SD 인문과학, SE 예술, SF 기타, SP 특허	413	56
합 계		2,642	1,462

[표 3] 분석대상 전자학술지 다운로드 횟수별 종수

기관 다운로드/종	B 대학	Y 대학	S 대학	P 대학	대학 전체	K 연구소	S 연구소	E 연구소	D 연구소	A 연구소	연구소 전체	합계
0회	250	210	143	470	102	454	968	1,193	1,350	785	276	95
1회 이상	2,175	2,291	2,377	2,031	2,520	1,987	1,345	1,111	936	1,564	2,231	2,547
11회 이상	1,615	1,764	2,065	1,245	2,253	1,252	705	390	325	733	1,614	2,301
101회 이상	677	778	1,332	501	1,606	488	249	88	41	198	780	1,681
1001회 이상	80	103	317	88	565	76	28	7	0	12	151	657
10001회 이상	0	2	17	3	42	1	0	0	0	0	4	61
(접근가능 학술지)	2,425	2,501	2,520	2,501	2,622	2,441	2,313	2,304	2,286	2,349	2,507	2,642

#### 1.4 분석 대상기관

분석 대상기관의 선정은 분석 대상 컨소시엄에 모두 가입되어 있는 기관으로서 이용량이 상위에 있는 기관을 선정코자 하였다. 그 결과 B대학, Y대학, S대학, P대학 등 4개 대학과, K연구소, S연구소, E연구소, D연구소, A연구소 등 5개 연구소를 선정하였는데 의료기관이나 기업연구소 등은 특정 컨소시엄에 편중되어 있어 동일 기준에 의한 비교가 되지 않으므로 분석 대상에서 제외하였다.

## 2. 전자학술지 이용분석

### 2.1 이용되지 않은 학술지

분석 대상 4개 대학과 5개 연구소에서 접근 가능한 전자학술지의 기관별 이용량 분석 결과를 표 3에 나타내었다. 이 표에는 각 학술지 타이틀의 아티클 다운로드 횟수 구간별로 기관별 학술지 종수를 표시하여 전자학술지의 개략적인 이용 정도를 나타내고 있다.

분석대상 학술지 2,642종을 대상기관이 1건도 다운로드 하지 않은 학술지는 95종인 3.6%이며, 이 중 Science Direct가 30종, Wiley가 65종이 포함되어있다. 이용되지 않는 학술지까지 번들 형식으로 제공되는 패키지 라이선싱의 문제 즉 키워 팔기와 관련한 문제가 지적되어 온 것과 관련하여 대학의 경우 기관별로 5-19%가 포함되어 있음을 확인했고, 연구소의 경우는 K 연구소는 18.5% 정도였으며, D 연구소는 59%로 나타나 전자저널 라이선싱을 보다 효율화 할 필요가 있음을 나타내고 있다.

### 2.2 1회 이상 다운로드 된 학술지

4개 대학은 전체적으로 접근이 가능한 2,622종 중에서 2,520종이 이용되어 96%의 이용률을 나타내고 있고, 연구소의 경우는 접근이 가능한 2,507종 중에서 2,237종이 이용되어 89%의 이용률을 나타내고 있다. 그러나 4개 기관의 1회 이상 평균

이용률에 있어 대학이 89%인 반면 5개 연구소의 경우 53%로서 상대적으로 낮다. 이용 대상자 수가 많고, 광범위한 주제를 연구하는 대학에서 다양한 학술지가 이용되는 것을 확인하였다.

### 2.3 1천회 이상 다운로드 된 학술지

1천회 이상 다운로드 된 학술지는 4개 대학 전체적으로는 22%인 565종이, 5개 연구소는 6%인 151종이었다. 9개 기관 전체적으로는 25%인 657종이 1천회 이상 다운로드 되었다. 역시 각 연구소가 대학에 비해 특정 주제 분야에 집중된 경향을 알 수 있다.

### 2.4 전자학술지 이용량과 JCR impact factor

분석대상 학술지 이용량 상위 1백종 내의 모든 학술지는 JCR 2006상에 impact factor가 부여되어 있었다. 그리고 JCR impact factor가 부여된 학술지의 평균적인 다운로드 횟수는 2,284회였으며, impact factor가 부여되지 않은 경우는 170회로 극단적인 차이를 보이고 있다. 이는 분명 impact factor와 이용량 간에 어느 정도 상관관계가 존재한다는 추정을 가능케 한다. 그러나 이용량 상위 5백종 중에서 JCR impact factor가 부여되지 않은 학술지가 28종이 포함되어 있어 절대적인 관계는 아님을 알 수 있다.

### 2.5 컨소시엄별 전자학술지 이용량

표 4는 분석대상 4개 컨소시엄의 학술지 종별로 각 기관의 평균 다운로드 횟수를 나타낸 것이다. 9개 기관 전체적으로는 Science 컨소시엄이 137,474건으로 단연 월등하며, ACS가 15,552건, ScienceDirect가 1,330건이며 Wiley가 634건으로 뒤를 이었다. 이는 표 1의 JCR 2006상의 impact factor 부여 학술지 비율이나 평균 impact factor 순위와 완전히 일치한다. 9개 기관 모두에서 이 순위는 일치하는 것으로 나타났다.

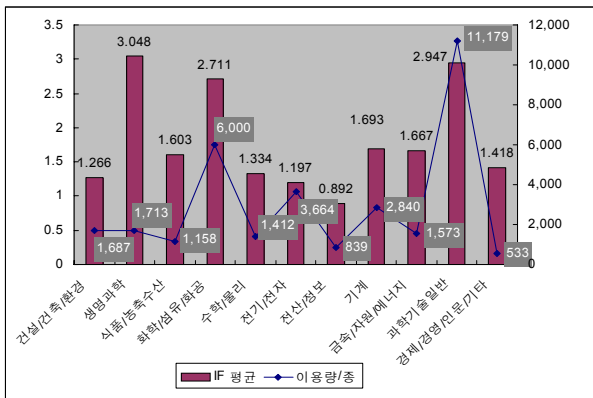
[표 4] 분석대상 전자학술지 종별 평균 다운로드 횟수

구분	ACS	Science	Science Direct	Wiley	계
B대학	1,085	7,418	173	59	8,734
Y대학	2,661	20,572	201	94	23,528
S대학	7,185	65,090	522	301	73,099
P대학	2,568	25,934	140	87	28,728
대학 계	13,499	119,014	1,036	541	134,089
K연구소	1,056	5,863	151	39	7,108
S연구소	568	9,501	64	29	10,162
E연구소	201	1,913	16	13	2,143
D연구소	37	233	9	5	284
A연구소	192	950	55	7	1,203
연구소 계	2,053	18,460	294	93	20,900
합계	15,552	137,474	1,330	634	154,990

### 3. 영향력 지수와 와 전자학술지 이용의 상관관계

#### 3.1 주제별 평균 impact factor와 전자학술지 이용량

11개 주제분야별 impact factor 평균과 전자학술지 이용량을 그림 1에 나타내었다. ‘생명과학’분야는 평균 impact factor가 3.0477로 가장 높음에도 불구하고 학술지 종당 이용량은 평균에 미치지 못했다. ‘과학기술 일반’ 분야는 impact factor 평균도 높고 종당 이용량은 가장 높은 11,179건을 나타냈다. 이는 분석 대상 학술지 중에서 impact factor 값과 이용량이 최상위에 있는 ‘Science’가 포함되고, 평균 이용량이 상위 1% 이내이며 impact factor가 상위에 속하는 ‘Nano Letters’와 같은 학술지가 이 분야에 포함되어 있지만 해당 학술지 주제분야 종수가 18종으로 매우 적기 때문에 풀이된다. ‘경영·경제·인문·기타’ 분야는 impact factor 평균에 비해 학술지 종당 평균 이용량은 상대적으로 매우 낮은 533건이었다.



▶▶ 그림 1. 주제별 impact factor 평균과 학술지 종당 평균 이용량

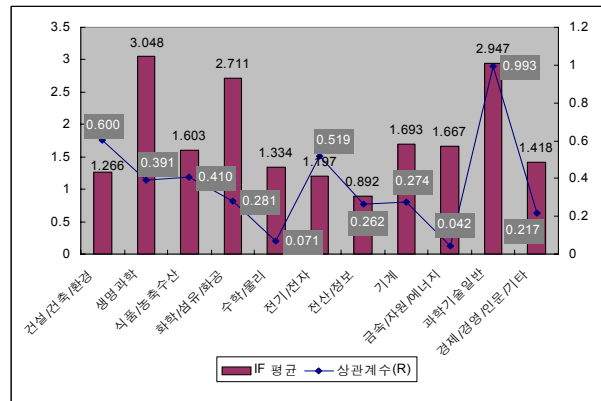
#### 3.2 학술지별 impact factor와 이용량의 상관관계

학술지별 impact factor와 이용량의 상관관계를 알아보기 위하여 Microsoft Excel을 이용한 상관관계 분석을 수행하였다. 분석결과 상관계수(R)은 0.37로 양의 상관관계가 있음을

확인했다. 그러나 impact factor가 부여되어 있다고 해서 전자 학술지의 이용량이 반드시 상대적으로 높다고 할 수는 없을 것으로 보인다. 그 이유는 분석대상 학술지 중 impact factor 값을 가지고 있는 7종의 경우 전혀 다운로드 되지 않았으며, 다운로드 건수가 100회 미만으로 매우 적은 학술지가 157종에 달하고 있기 때문이다.

#### 3.3 주제분야별 impact factor와 이용량의 상관관계

그림 2에 주제분야별로 impact factor와 이용량의 상관계수(R)을 나타내었다. 또한 impact factor 평균을 막대로 표시하여 대조할 수 있도록 하였다.



▶▶ 그림 2. 주제분야별 impact factor와 이용량의 상관관계

상관관계가 가장 높은 주제분야는 ‘과학기술일반’ 분야로 상관계수(R)이 0.993에 달하였다. 그러나 3.1에서 설명한대로 18종에 불과한 학술지의 impact factor와 이용량에 의해 도출되었으므로 크게 신뢰할 수는 없다. ‘건설·건축·환경’ 분야의 경우 상관계수(R)이 0.6으로 높았고 전기전자 분야도 0.5를 상회하여 높은 상관관계가 있음을 나타내고 있다.

그림 1의 이용량 선 그래프와 그림 2의 상관계수 선 그래프는 일견 닮아 있는 듯하나, ‘과학기술일반’ 분야를 제외하면 상당한 차이를 보이고 있다. 특히 ‘건설·건축·환경’ 분야와 ‘화학·섬유·항공’ 분야의 경우 주제별 순위의 차이가 매우 크게 나타난다.

## IV. 결론

JCR impact factor가 분석 대상인 4개 컨소시엄의 전자 학술지 이용량과 어느 정도 상관관계가 있는 것을 확인하였다. 그러나 주제별로 상관관계의 편차가 매우 크고, JCR impact factor의 유무나 높고 낮음이 이용량을 판단할 수 있는 척도가 되지 못한다는 것도 확인하였다.

많은 기관이 SCI 학술지 여부를 가지고 학술지의 품질을 판단하고, 정보자원 개발에도 활용하고 있다. 또한 연구자의 연구능력이나 기관을 평가할 때에도 유력한 척도로 사용되고 있다. 하지만 학술지의 품질을 판단하거나 이용량을 예측하기 위한 다른 대안이 없을 때와는 방법이 달라져야 할 것이다. 오래 전부터 학술지 이용량의 측정과 관련하여 어려움을 겪어왔지만, 이제 전자학술지가 연구자에게 보편화되었으므로 학술지 이용량을 측정하기가 기술적으로 용이해졌다. 어떤 학술지 또는 어떤 논문을 누가 언제 이용했는지를 쉽게 측정할 수 있는 것이다. 그러나 이용통계를 상업 출판사에 의존해야 하며, 통계 데이터가 표준화 되어있지 않아 수집과 분석에 있어서의 어려움이 아직 존재한다. 이러한 문제를 해결해 나가면서 Joan 등이 시도한 usage impact factor 개발이 필요하며 이를 위한 연구가 계속되어야 할 것이다.

#### ■ 참고 문헌 ■

- [1] Research Information Network. "UK Scholarly Journals: 2006 Baseline Report: an Evidence-based Analysis of Data Concerning Scholarly Journal Publishing," [http://www.rin.ac.uk/files/UK%20scholarly%20Journals%202006%20Baseline%20Report.pdf] (Accessed October 6, 2007)
- [2] Bollen, Johan; Van de Sompel, Herbert, "Usage Impact Factor: the effects of sample characteristics on usage-based impact metrics", eprint arXiv:cs/0610154
- [3] Blečić, Deborah D.; Fiscella, Joan B.; Wiberley, Stephen E., Jr. "The Measurement of Use of Web-based Information Resources: An Early Look at Vendor-supplied Data", *College & Research Libraries*, Vol. 62, No. 5, pp. 434-53
- [4] King, Donald W., Tenopir, Carol Michael Clarke, "Measuring Total Reading of Journal Articles", *D-Lib Magazine*, Vol. 31 No. 10, 2006
- [5] JISC, "NESLi2 Analysis of Usage Statistics: Summary Report," [http://www.jisc.ac.uk/uploaded\_documents/nesli2\_usstudy.pdf] (Accessed October 5, 2007)