

완벽주의 지수 PI의 개량을 통한 유력 학술지와 대량생산 학술지의 구분

Improving the Perfectionism Index to Identify Influential Journals versus Mass Producers

이재윤 (Jae Yun Lee)*

초 록

최근 제안된 완벽주의 지수 PI는 연구자를 유력자와 대량생산자로 구분하는 지표이다. 이 연구에서는 PI를 개량한 새로운 지표인 준완벽주의 지수 NPI를 제안하였다. NPI는 특히 발행시기 등을 고려하지 않고 저인용논문에 무조건 획일적인 기준으로 패널티를 부과하던 PI의 방식을 개선하는 보완 지수이다. NPI에서는 꼬리 보상 영역에 인용빈도 곡선을 고려하면서 패널티를 부과함으로써, h-지수의 향상이 오히려 영향력 지표에 불리하게 작용하는 것을 방지한다. 이렇게 개발된 NPI를 Web of Science 문헌정보학 관련 분야 학술지에 시험 적용해본 결과 h-지수와 평균 인용횟수로는 불가능했던, 유력 학술지와 대량생산 학술지의 구분을 성공적으로 수행할 수 있었다.

ABSTRACT

The Perfectionism Index (PI) is an indicator that is recently proposed to distinguish influential researchers from mass producers. In this study, Near Perfectionism Index (NPI), an improved indicator of Perfectionism Index, can be a solution to the problem of PI that indiscriminately gives a penalty to all low-cited papers regardless of publishing time or other issues. NPI improved the method to give a penalty to tail complement area considering the citation distribution curve. It prevents the improvement of the h-index from adversely affecting the researcher's influence indicator. This study uses NPI to evaluate information and library science journals in Web of Science database. It successfully distinguishes between influential journals and mass producers unlike journal h-index or average citation frequency which could not differentiate influentials from mass producers.

키워드: 인용 지수, 완벽주의 지수, 준완벽주의 지수, 학술지 평가, 연구자 평가, h-지수, 유력 학술지, 거대 학술지

citation indicator, perfectionism index, near perfectionism index, journal evaluation, researcher evaluation, h-index, influential journal, mega journal

* 명지대학교 문헌정보학과 교수(memexlee@mju.ac.kr)

■ 논문접수일자: 2019년 5월 20일 ■ 최초심사일자: 2019년 6월 24일 ■ 게재확정일자: 2019년 6월 24일
■ 정보관리학회지, 36(2), 201-222, 2019. [http://dx.doi.org/10.3743/KOSIM.2019.36.2.201]

1. 서론

연구자를 평가하는 인용 지수인 h-지수(Hirsch, 2005)는 한국학술지인용색인(Korea Citation Index: 이하 KCI로 표기)에서 연구자 성과 지표로 제공할 정도로 국내에서도 널리 쓰이고 있다. h-지수는 연구자 평가뿐만 아니라 기관 평가(박지연, 김정은, 민윤경, 2010), 학술지 평가(김관준, 이재운, 2010)에도 활용하는 방안이 제안된 바 있다.

최근 Sidiropoulos, Katsaros, Manolopoulos (2015)는 h-지수를 산출하는 문헌 인용횟수 그래프를 활용해서 연구자를 '유력자(influentials)'와 '대량생산자(mass producers)'로 구분하는 방안으로 '완벽주의 지수(Perfectionism Index: 이하 PI로 표기)'를 제시하였다. 이들의 구분에 의하면 유력자는 완벽주의 성향이 커서 연구의 양보다는 연구의 질을 우선시하는 연구자이다. 결함이 없는 완벽한 논문을 추구하므로 유력자가 발표한 논문의 인용 영향력은 대부분 높기 마련이다. 반면에 대량생산자는 연구의 질보다는 연구의 양을 우선시하는 연구자로서 보통이거나 낮은 영향력을 가진 논문을 다수 발표하게 된다. 이들이 h-지수에 만족하지 않고 새로운 지수를 제안한 이유는 h-지수만 보면 연구자의 논문별 인용횟수 분포에 담긴 추가 정보를 놓치기 때문이다. h-지수로는 소수의 높은 영향력이 큰 논문 위주로 발표하는 연구자와 다수의 보통 이하 영향력의 논문을 주로 발표하는 연구자를 구분하기 어렵다. 극단적인 예를 들자면, 100회 인용된 5편의 논문을 발표한 연구자와 5번 인용된 100편의 논문을 가진 연구자가 동일하게 h-지수 5로 평가되므로 이들의

차이를 알아볼 수 없는 것이다.

이 연구에서는 연구자 개인보다는 학술지의 특성을 파악하는 용도로 완벽주의 지수 PI와 이를 개량한 지수를 사용할 것을 제안하고자 한다. 연구자가 아닌 학술지 차원에서 유력자와 대량생산자, 즉 유력 학술지와 대량생산 학술지를 구분하는 것이 필요한 이유는 다음과 같다.

첫째, 더욱 보편화되는 학술지 등급제와 영향력 지수 기준 학술지 평가의 확산 때문에 게재 논문이 대부분 일정 수준 이상임을 알려주는 새로운 지표의 필요성이 커졌다. 원론적으로 연구성과 평가는 논문 자체의 우수성을 기준으로 이루어져야 한다. 그렇지만 우리나라에 서뿐만 아니라 세계적으로도 논문이 게재된 학술지의 명성이나 등급을 기준으로 논문을 평가하는 바람직하지 않은 현상이 점차 확산되면서 이런 현상을 경계하는 전문가의 목소리도 커지고 있다(유소영, 이재운, 정은경, 이보람, 2015). 이제는 국내 논문 중에서 한국연구재단 등재/등재후보 학술지에 게재된 논문이 아니면 연구성과로 제대로 인정받지 못하는 것이 일반화되었고, 2015년부터는 KCI에 '우수등재 학술지' 등급이 신설되면서 2018년까지 67종의 학술지가 우수등재 학술지로 선정되었다. 앞으로 몇 년 뒤에는 우수등재/등재/등재후보/미등재의 4등급 중에서 어디에 속하는 학술지에 논문을 게재했느냐가 연구성과 평가에 중요해질 수 있다. 또한 KCI 등재지 평가에서 평균 인용횟수에 해당하는 영향력 지수도 점점 더 중요하게 고려되고 있다. 그런데 설사 평균 인용횟수가 높더라도 소수의 고인용 논문이 평균을 끌어올리는 경우가 많고 상당수의 저인용 논문이 학

술지의 높은 영향력 지수에 무임승차하게 된다 (Baum, 2011). 따라서 높은 등급의 학술지가 게재된 논문의 질적 수준을 그만큼 엄격하게 유지하는가를 검증해야 될 필요성도 커졌다.

둘째, 거대 학술지(mega journal)가 급속히 확산되면서 유력 학술지와 대량생산 학술지를 구분해야 할 필요성이 커졌다. 거대 학술지는 광범위한 학문 분야를 포괄하면서 오픈 액세스 방식으로 대량의 논문을 발행하되, 심사 과정에서 연구 방법의 타당성만 검토하고 주제의 학술적 중요도는 상관하지 않고 게재 후 독자에게 판단을 맡기는 학술지이다(Björk, 2015). *PLOS One*이 가장 대표적인 사례로서 거대 학술지를 ‘*PLOS One*식 학술지(*PLOS One*-like journal)’라고 부르기도 한다. *PLOS One*의 경우는 2006년 창간년도에 1,239편의 논문을 게재한 이후 급성장하여 가장 많았던 2013년에는 한 해 동안 31,404편을 게재하기에 이르렀다 (Spezi, Wakeling, Pinfield, Creaser, Fry, & Willett, 2017). 이후 게재 논문 수가 다소 감소하였으나 2017년에도 연구논문만 19,983편이 게재되었고, 리뷰까지 포함하면 2만 건이 넘었다(Web of Science 검색 결과). *PLOS One*은 공식적으로 밝히기를, 기술적으로 타당하다고만 인정되면 내용의 참신성이나 연구 결과의 성공 여부와 상관없이 모든 논문을 게재하며 학술적 중요도에 대해서는 발간 후에 독자들이 판단하게 하는 정책을 유지하고 있다고 한다(*PLOS ONE journal information*). 이제는 *PLOS One* 처럼 논문 대량생산에 치중하는 거대 학술지와, 중요한 논문만 엄격하게 선별하여 게재하는 유력 학술지를 구분할 수 있는 방법이 더욱 필요해졌다. 더군다나 국내에서도 공공 연구성과물

의 오픈 액세스 출판에 대한 정책적 검토가 강화되면서(민윤경, 차미경, 2017) 오픈 액세스 학술지의 특성을 파악할 수 있는 지수의 개발이 시급해지고 있다.

셋째, 완벽주의 지수 PI 산출 과정에서는 h-지수값보다 적게 인용된 논문에 페널티를 부과하게 되는데, 개인 연구자를 평가할 경우에는 논문마다 가지는 인용될 기회의 차이가 큰 반면에 학술지를 평가할 경우에는 논문마다 인용될 기회가 더 균등하게 부여될 수 있기 때문에 연구자보다 학술지를 평가하는 것이 더욱 타당하다. 개인 연구자의 h-지수를 산출할 때에는 평생의 연구 논문을 모두 포함하므로 각 논문마다 인용될 시간적인 기회가 매우 불균등한 반면에, 학술지의 h-지수 산출에는 몇 년 이내의 일정 기간 동안 발표된 논문만을 포함하므로 인용될 시간적인 기회가 상대적으로 더 균등하다. PI를 산출하는 과정에서 h-지수보다 낮은 인용횟수를 가진 논문은 모두 페널티를 부과받아서 지수값을 낮추는 결과를 가져온다. 그런데 어떤 논문의 인용횟수가 낮거나 0에 머무는 것은 논문의 내용이 미흡하기 때문일 수도 있으나 다른 이유도 얼마든지 있다. 당장 논문이 발표된 직후에는 인용될 기회가 없었으므로 인용횟수가 낮을 수밖에 없다. 또한 학계의 관심이 적은 주제를 다루어서 인용되지 못할 수도 있으며, 때로는 발표 후 몇 년에서 수십 년까지 지난 후에야 뒤늦게 주목받고 활발하게 인용되기에 소위 ‘잠자는 미녀(sleeping beauty)’(Van Raan, 2004)로 불리는 논문도 존재한다. 이렇게 다양한 이유가 가능성에도 불구하고 개인 연구자가 연구경력 전체에서 발표한 모든 논문이 동일한 인용 기회를 가진 것으로 전제하고

서 특정 수준보다 인용횟수가 낮은 논문은 모두 '완벽하지 않은' 연구 성과물이라고 치부하는 것은 불공정한 판단이다. 아무리 저자가 완벽을 추구한 뛰어난 논문이라고 하더라도 적어도 발표 직후에는 인용횟수가 모두 0일 수밖에 없기 때문이다. 따라서 연구활동이 왕성하여 최근 1~2년 사이에 다수의 논문을 발표한 연구자는 손해를 볼 수밖에 없다. 한편 학술지를 대상으로 h-지수를 산출할 때에는 인용횟수를 측정하는 대상 논문의 발행 시기 범위를 1년(Braun, Glänzel, & Schubert, 2006), 3년(김관준, 이재운, 2010; Scimago Journal & Country Rank), 5년(Google Scholar: Harzing & van der Wal, 2009) 정도로 일정 기간 이내로 제한하는 것이 보통이다. 따라서 개인 연구자의 평생 업적으로 평가할 경우에 비해 학술지의 경우는 논문마다 인용될 기회의 차이가 덜하게 되므로 인용횟수가 일정 수준 미만인 논문에 대해서는 패널티를 부과하는 것이 상대적으로 더 정당성을 가지게 된다. 다만, 그렇다고 하더라도 발행된 직후의 논문이 인용될 기회가 가장 적은 것은 사실이므로 이를 고려할 개선 방안이 필요하다. 이에 대해서는 뒤에서 더 상세히 다루었다.

이 연구에서는 개인 연구자를 유력 연구자와 대량생산자로 구분하기 위한 완벽주의 지수 PI를 응용하여 유력 학술지와 대량생산 학술지를 구분하여 파악하기 위한 방안을 개발하고자 한다. 특히 발행시기를 고려하지 않고 저인용논문에 무조건 획일적인 기준으로 패널티를 부과하던 방식을 개선하는 보완책을 제안한다. 이를 위해서 2장에서는 완벽주의 지수 PI에 대해서 설명하고, 3장에서는 PI를 개선한 새로운 지

수를 제안한 후, 4장에서 Web of Science 문헌정보학 분야 학술지에 대해 PI와 새로운 지수를 적용해본 결과를 설명하였다.

2. 완벽주의 지수 PI의 개념과 문제점

2.1 완벽주의 지수 PI의 개념

연구자가 논문을 발표하는 행태는 소속 기관이나 국가의 고등교육 및 연구개발 정책에 좌우되기도 하지만, 학문 분야나 개인의 성향에 따라서도 달라질 수 있다. 오래 전 Cole과 Cole(1967)은 논문 발표 행태에 따라서 연구자의 유형을 4가지로 구분하면서 소수의 영향력이 높은 논문을 제한적으로 발표하는 완벽주의자(perfectionist)와 영향력이 낮은 논문을 다수 발표하는 대량생산자(mass producer)를 대비하였는데, Costas와 Bordons(2008)는 유사한 4가지 유형으로 구분하면서 완벽주의자와 같은 유형을 'selective', 대량생산자와 같은 유형을 'big producers'라고 불렀다. 이런 구분은 Rosenberg(2014)를 비롯해서 최근까지도 계속 답습되고 있다.

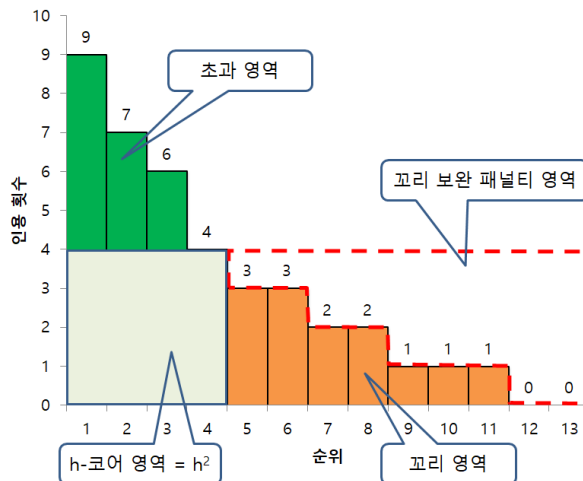
Sidiropoulos, Katsaros, Manolopoulos(2015)는 선행 연구에서 구분한 방식을 따르면서 많은 논문을 발표했지만 대부분이 상대적으로 인용횟수가 낮은 연구자는 '대량생산자(mass producer)'라고 칭하고, 대부분의 발표 논문이 고인용 논문인 연구자는 '유력자(influential)'라고 구분하였다. 연구 결과를 발표하는 행태를 보면 유력자는 완성도가 덜한 상태의 논문을 발표하기를 꺼

려하는 성향을 가지므로 발표 논문 수는 적은 반면 발표한 논문의 영향력은 큰 편이기 마련이고, 반대 성향의 연구자들은 아이디어를 단계적으로 발전시키면서 보통 수준의 영향력을 가진 논문을 더 빈번하게 발표하는 경향이 있다는 것이다.

유력자와 대량생산자를 구분하기 위해서 Sidiropoulos, Katsaros, Manolopoulos(2015)는 h-지수를 산출하는 인용 곡선(citation curve)으로부터 유력자와 완벽주의 지수(Perfectionism Index) PI를 도출하였다. 유력자는 자신의 h-지수값보다 그다지 많지 않은 갯수의 논문을 발표했으나 대부분이 높은 인용횟수를 얻은 연구자로서 PI 값이 큰 양수가 된다. 반면에, 대량생산자는 자신의 h-지수 값보다 훨씬 많은 분량의 논문을 발표했지만 대부분이 h-지수 값보다 꽤 낮은 인용횟수를 얻은 연구자로서 PI 값이 낮은 음수가 된다.

h-지수를 산출하는 과정에서 연구자의 논문

별 인용횟수 분포 구간은 <그림 1>의 막대그래프와 같이 세 영역으로 분할된다. 첫째는 h회 이상 인용된 h개의 논문으로 구성되는 정사각형 영역으로서 'h-코어'라고 부른다(Ye & Rousseau, 2010). h-코어 영역의 넓이는 h의 제곱에 해당한다. h-코어 영역의 위쪽과 오른쪽에 위치한 나머지 두 영역은 각각 초과(excess) 영역과 꼬리(tail) 영역이라고 부른다. Sidiropoulos, Katsaros, Manolopoulos(2015)는 완벽주의 성향을 가진 유력 연구자라면 자신의 모든 논문이 일정 수준 이상의 영향력을 가지도록 완성도를 높여 발표할 것이라는 가정을 하였다. 그들은 기준이 되는 '일정 수준'을 연구자의 h-지수 값이라고 전제한 후, h회보다 적게 인용된 꼬리 영역의 논문들은 완벽주의 성향이 덜 발휘된 결과이기 때문에 h에 못미치는 정도만큼 패널티를 적용해야 한다고 주장하고 이를 꼬리 보완 패널티 영역이라고 명명하였다.



<그림 1> 가상 연구자 A의 논문별 인용횟수 그래프에서 h-코어 영역, 초과 영역, 꼬리 영역, 꼬리 보완 패널티 영역 구분

예를 들어 <그림 1>과 같이 h-지수가 4인 가상의 연구자 A가 발표한 13편의 논문이 각각 얻은 인용횟수가 {9,7,6,4,3,3,2,2,1,1,1,0,0}이라면, 4회 이상 인용된 상위 4편의 논문이 이루는 정사각형 영역이 h-코어 영역이고 면적은 $4 \times 4 = 16$ 이다. h-코어를 구성하는 상위 4편의 논문이 4회를 초과하여 인용된 횟수가 이루는 영역이 초과 영역(C_E 로 표기)이고 크기는 $(9+7+6+4) - (4 \times 4) = 10$ 이다. h-코어 오른쪽에 5위 이하의 논문이 인용된 횟수가 이루는 영역은 꼬리 영역(C_T 로 표기)이고 크기는 $3+3+2+2+1+1+1 = 13$ 이다. 마지막으로 꼬리에 속한 하위 9편의 논문들이 h보다 적게 인용된 만큼의 차이가 난 부분이 꼬리 보완 페널티 영역(C_{TC} 로 표기)이고 크기는 $(9 \times 4) - 13 = 23$ 이다. 완벽주의 지수 PI는 다음 공식과 같이 h-코어 영역과 초과 영역의 넓이를 더한 후 꼬리 보완 페널티 영역의 넓이만큼 빼서 산출한다(Sidiropoulos, Katsaros, & Manolopoulos, 2015).

$$\text{완벽주의 지수 } PI = h^2 + C_E - C_{TC}$$

꼬리 보완 페널티 영역의 넓이는 다음과 같이 h에 꼬리 영역 논문 수를 곱한 직사각형의 넓이에서 꼬리 영역의 넓이 C_T 를 빼서 산출한다.

$$\text{꼬리 보완 페널티 영역 } C_{TC} = h \times (P - h) - C_T$$

가상 연구자 A의 경우 PI는 $16+10-23 = 3$ 이 된다. 원래 Sidiropoulos, Katsaros, Manolopoulos (2015)는 PI를 제안하면서 공식의 세 항 각각의 반영 비율을 조절할 수 있는 파라미터를 설정하였으나, 시험 적용해보면서 파라미터의 조

정이 필요없다고 판단하여 실제로는 사용하지 않았다.

한편 가상 연구자 A의 PI 값 3은 논문 13편의 총 인용횟수 39에서 꼬리 영역 논문 수 9에 h-지수값 4를 곱한 결과인 36을 뺀 것과 같다. 따라서 총 인용횟수를 C 라고 하고 꼬리 영역 논문 수를 P_T 라고 할 때 h-지수가 h인 연구자의 PI를 산출하는 공식은 더 단순하게 다음과 같이 표현할 수도 있다.

$$\text{완벽주의 지수 } PI = C - h \times P_T$$

위 공식에 따르면 총 인용횟수가 일정할 때 어떤 연구자의 PI가 가질 수 있는 최댓값은 꼬리 영역의 논문이 모두 h회만큼씩 인용되었을 경우 즉 꼬리 보완 페널티 영역이 0일 때 달성되고, 그때 PI 값은 h-코어에 속한 논문의 인용횟수 합계와 같다. 반대로 어떤 연구자의 PI가 가질 수 있는 최솟값은 꼬리 영역의 논문이 모두 인용횟수가 0일 때 달성되고, 그때 PI 값은 미인용된 논문 수에 h를 곱한 만큼을 총 인용횟수에서 뺀 값이다. 한편 논문 수가 많지 않은 연구자는 h-지수도 작고 초과 영역이나 페널티 영역이 커질 여지도 별로 없으므로 PI도 중립 값인 0에 가깝게 된다. 논문 수가 어느 정도 이상이 되어야 PI가 높은 양수가 되거나 낮은 음수가 될 수 있어서 유력자나 대량생산자임을 파악할 수 있다. 이런 현상은 Sidiropoulos, Katsaros, Manolopoulos(2015)가 실제 데이터에 적용했을 때에도 나타났다.

Microsoft Academic Search의 컴퓨터 과학 분야 연구자 데이터를 활용해서 Sidiropoulos, Katsaros, Manolopoulos(2015)가 분석해본 결

과, 유력자로 파악되는 연구자들은 산업계의 중요한 장벽을 돌파한 방안의 창안자(inventors)로서 상대적으로 적은 논문을 발표한 경우이거나, 지지벡터기계를 개발한 Vapnik나 퍼지 논리를 제안한 Zadeh와 같이 독자적으로 이론이나 기법을 개발하고 지속적으로 발전시킨 창시자들이 대부분이었다. 이런 유력자들은 학술적인 측면에서 학계에 큰 영향을 끼친 연구자들이라고 할 수 있다. 반면에 대량생산자로 파악되는 연구자들은 대부분 많은 제자를 양성하면서 꾸준히 멘토 역할을 수행하여 제자들과 공저로 다수의 논문을 발표한 경우였다. 제자를 학자로 양성하는 과정에서 생산된 논문이다 보니 학술적인 완성도는 상대적으로 높지 못한 경우가 흔했다. 멘토 역할을 수행한 대량생산자들은 인적인 측면에서 학계에 크게 기여한 연구자들이다.

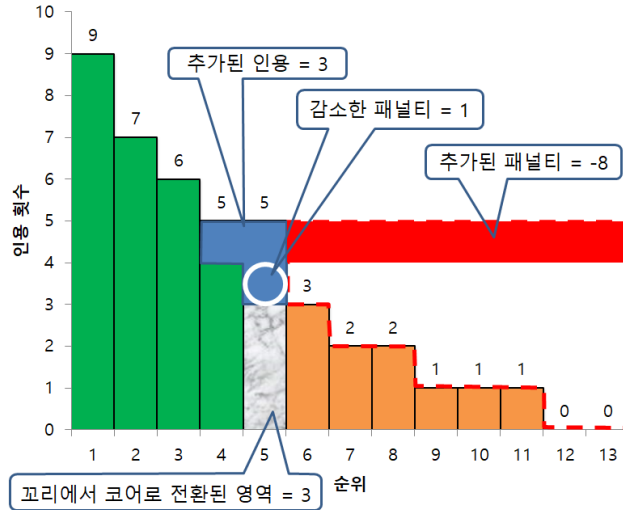
2.2 완벽주의 지수 PI의 한계

완벽주의 지수 PI로 평가할 경우에 논문 수가 상대적으로 적은 연구자는 값이 중립적인 0으로부터 크게 벗어나기 어렵기 때문에 유력자나 대량생산자가 될 수 없다. 이 때문에 PI로 구분하게 되는 연구자의 유형은 높은 양의 값을 가지는 유력자와 낮은 양의 값을 가지는 대량생산자의 두 종류 이외에, 0에 가까운 값을 가지는 연구자 유형이 추가된다. PI가 0에 가까운 연구자들은 대부분이 Cole과 Cole(1967)이 언급한 4가지 유형 중 논문 수도 적고 영향력도 낮은 잠잠한(silent) 연구자 유형이라고 할 수 있다. 같은 경우를 Costas와 Bordons(2008)는 'low producer' 유형이라고 구분하였다. 그러나

때로는 Cole과 Cole(1967)이 언급한 '풍성한 연구자' 유형에 속해서 평균 인용횟수와 논문 수가 모두 많은 연구자라 할지라도 PI가 0에 가깝게 될 수도 있다. 이들은 높은 h-지수와 넓은 초과 영역을 확보하지만 때로는 높은 h-지수 때문에 오히려 많은 꼬리 영역의 논문에 큰 패널티를 받게 될 수도 있는 것이다.

특히 꼬리 보완 패널티 영역의 크기를 좌우하는 것이 h-지수의 크기이기 때문에 같은 논문 수로 인용을 더 얻어서 h-지수와 평균 인용횟수는 증가했는데도 불구하고 PI는 감소하는 경우가 발생할 수 있다. Cole과 Cole(1967), 그리고 Costas와 Bordons(2008)의 기준에서는 평균 인용횟수의 상승이 연구의 질적 수준 상승을 의미하는데, 논문 수는 같고 질적 수준은 상승했는데도 불구하고 완벽주의 성향을 의미하는 PI는 오히려 감소하는 경우도 발생하는 것이다. 앞의 가상 연구자 A의 경우를 활용해서 이런 상황을 구체적으로 살펴보기로 한다.

가상의 연구자 A가 발표한 13편의 논문이 각각 얻은 인용횟수가 {9,7,6,4,3,3,2,2,1,1,1,0,0}이었는데, 여기서 4번째 논문과 5번째 논문이 각각 1회와 2회 더 인용되어 인용횟수 목록이 {9,7,6,5,5,3,2,2,1,1,1,0,0}으로 되었다고 가정한다. 그러면 h-지수는 4에서 5로, 평균 인용횟수는 3.00에서 3.23으로 상승한다. 이처럼 논문수는 같아도 총 인용횟수, 평균 인용횟수, h-지수가 모두 상승하여 질적인 연구성과는 더 높아진 것으로 측정된다. 그런데 h-지수가 1만큼 증가했기 때문에 꼬리 영역의 논문 수가 9개에서 8개로 감소했지만, 꼬리 보완 패널티 영역을 정하는 기준이 4에서 5로 높아져서 꼬리 영역 논문의 8개의 패널티가 모두 1씩 증가하여 <그림 2>와



<그림 2> 가상 연구자 A의 4번째와 5번째 순위 논문의 인용횟수가 증가하여 h-지수가 4에서 5로 상승하였을 경우의 각 영역의 변화

같이 꼬리 보완 패널티가 23에서 30으로 대폭 증가하게 된다. 결과적으로 평균 인용횟수와 h-지수로 측정된 질적인 연구성고가 향상되었음에도 불구하고 PI는 <표 1>과 같이 오히려 3에서 2로 감소하는 결과가 나타난다.

어떤 논문이 더 인용되거나 덜 인용되는 것에는 논문의 질적 수준 이외에도 다양한 원인

이 존재한다. 다루고 있는 주제가 '빅데이터'와 같이 시류를 타는지 여부, 학문 분야별 관행 차이, 연구자 집단의 규모가 많은 분야와 소외 분야의 차이 등에 의해서도 인용횟수가 차이날 수 있다. 이보다 더 뚜렷한 인용횟수 차이의 원인은 발행 후 경과된 시간이다. 지난 달에 발행된 최신 논문과 5년 전 발행된 논문, 10년 전

<표 1> 가상 연구자 A의 4번째와 5번째 순위 논문의 인용횟수가 증가하여 h-지수가 4에서 5로 상승하였을 경우(A1)의 변화

구분	A의 경우	A1의 경우
논문수 (P)	13	13
총 인용횟수 (C)	39	42
평균 인용횟수 (C/P)	3.00	3.23
h-지수	4	5
h-코어	16	25
초과 영역 (C_E)	10	7
꼬리 영역 (C_T)	13	10
꼬리 보완 패널티 영역 (C_{TC})	23	30
완벽주의 지수 PI	3	2

발행된 논문이 각각 얻는 인용될 기회는 현격하게 다르다. 심지어는 발행월에 따라서도 통계적으로 인용횟수가 달라진다는 보고도 있다 (Donner, 2018). 최근까지도 지속적으로 논문을 활발히 발표하고 있는 연구자는 과거 논문에 의해 높은 h-지수를 얻더라도 최근에 발표한 여러 논문이 시간이 부족하여 h에 미치지 못하는 적은 인용을 얻은 바람에 꼬리 보완 페널티가 커지게 된다. PI를 이상적으로 산출하려면 연구자의 연구 활동을 중지시키고 몇 년이 경과하여 대부분의 논문이 인용을 얻을 기회를 충분히 가진 이후에야 공정한 평가가 가능한 것이다. 현실적으로는 최근 몇 년간의 논문은 제외하고 지수를 산출할 수밖에 없다.

이상과 같이 완벽주의 지수 PI는 꼬리 영역 논문에 페널티를 부과하는 기준이 h-지수로 고정되어 있기 때문에 h-지수의 향상이 오히려 부정적으로 작용하는 경우가 발생할 수 있다. 또한 현실에서는 인용될 시간의 차이, 주제의 특성 등에 의해서 일부 논문은 인용을 적게 받을 수밖에 없음에도 불구하고 이를 연구자의 책임으로 치부하여 h보다 작은 인용을 받은 모든 논문에 동일한 페널티를 부과하는 것은 다소 비현실적이다.

3. PI의 개선 방안

논문마다 가지는 인용될 가능성이 균등하게 같을 수는 없다. 학술논문의 인용 분포는 멱함수 분포, 또는 로그정규분포를 보이면서 소수의 고인용 논문과 다수의 저인용 논문이 존재하는 것으로 알려져 있다(Brzezinski, 2015; Redner,

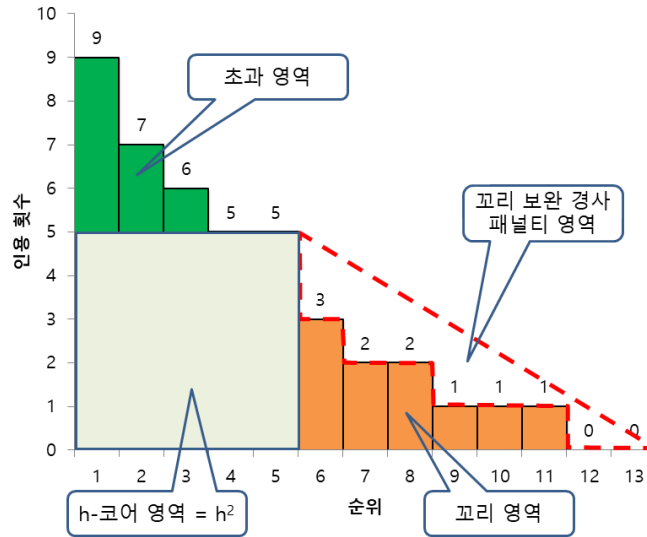
1998). 따라서 이 연구에서는 h보다 적게 인용된 논문에 차등화된 페널티를 부과하도록 PI를 개선해보았다.

새로운 지수는 준완벽주의 지수(Near Perfectionism Index) NPI로 명명하였다. 준완벽주의 지수 NPI는 초과 영역과 h-코어 영역은 PI와 동일하게 적용한다. 다만 꼬리 보완 페널티 영역의 산출 기준으로 PI에서는 모든 꼬리 영역 논문이 일괄적으로 h만큼 인용횟수를 가질 것을 요구하는 것에 비해서, NPI에서는 꼬리 영역 논문의 인용횟수가 h부터 0까지 감소하는 것을 전제로 한다. 현실적으로 인용횟수 분포는 멱함수 분포, 또는 로그정규분포를 보이므로 상위권 논문들 사이에서는 인용횟수의 감소가 가파르고 하위권 논문들 사이에서는 인용횟수의 감소가 완만하다. 순위 하락에 따른 인용횟수의 감소는 피할 수 없으나 감소의 정도가 선형적인 경우를 이상적인 목표 수준으로 설정하고, 목표 수준에 가까워서 인용의 빈익빈 부익부 현상이 완화되는 정도만큼 페널티가 감소하도록 지수를 설계한다. PI에서 꼬리 보완 페널티 영역의 역할을, NPI에서는 꼬리 보완 경사 페널티(tail complement slope penalty) 영역이 <그림 3>과 같이 수행하도록 변경한다.

꼬리 보완 경사 페널티 영역 C_{TCS} 는 다음 공식과 같이 꼬리 영역 논문 수에 h-지수를 곱한 결과의 절반으로부터 꼬리 영역 인용횟수를 뺀 값이다.

$$\text{꼬리 보완 경사 페널티 영역 } C_{TCS} = h \times \frac{P-h}{2} - C_T$$

준완벽주의 지수 NPI는 다음 공식과 같이 h-코어 영역과 초과 영역의 넓이를 더한 후 꼬리

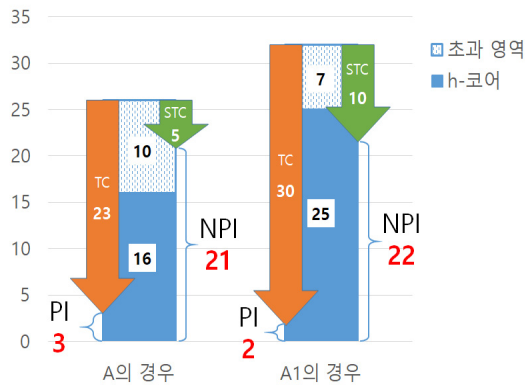


〈그림 3〉 준완벽주의 지수 NPI를 산출하기 위한 h-코어 영역, 초과 영역, 꼬리 영역, 꼬리 보완 경사 패널티 영역 구분

보완 경사 패널티 영역의 넓이만큼 빼서 산출하며, 이는 총 인용횟수 C에서 꼬리영역 논문수 P_T 와 h-지수를 곱한 값의 절반을 뺀 것과 같다.

$$\begin{aligned} \text{준완벽주의 지수 } NPI &= h^2 + C_E - C_{TCS} \\ &= C - \frac{h \times P_T}{2} \end{aligned}$$

앞서 살펴보았던 사례에서 가상의 연구자 A의 h-지수와 평균 인용횟수가 모두 상승했음에도 불구하고 PI는 오히려 감소하였다. 이는 〈그림 4〉에서 보듯이 h-코어의 증가가 꼬리 보완 패널티 영역을 크게 확장시키고 또한 초과 영역의 일부 축소로 이어졌기 때문이다. 그런데 NPI



〈그림 4〉 가상 연구자 A의 4번째와 5번째 순위 논문의 인용횟수가 증가하여 h-지수가 4에서 5로 상승하였을 경우(A1)에 각 영역이 PI와 NPI 산출에 기여하는 정도

를 기준으로 보면 인용이 추가된 이후에 지수는 21에서 22로 약간 상승하게 된다. 이는 h-코어의 증가로 인한 꼬리 보완 경사 패널티 영역의 확장과 초과 영역의 축소가, 늘어난 h-코어 영역을 상쇄할만큼 크지 않기 때문이다. 결과적으로 연구자 A의 논문 수는 그대로이고 인용횟수가 증가하여 평균 인용횟수와 h-지수가 소폭 상승했을 때 완벽주의 지수 PI는 감소했지만, 준완벽주의 지수 NPI는 소폭 상승한 것으로 나타났다. 이는 준완벽주의 지수 산출 과정에서 꼬리 영역의 논문이 무리하게 h-지수만큼 인용될 것을 요구하지 않고 합리적으로 차등 패널티를 부여했기 때문이다.

다음 장에서는 PI와 NPI를 실제 학술지 데이터에 적용하는 방법과 지수 산출 결과를 분석해보기로 한다.

4. Web of Science 학술지 평가 실험

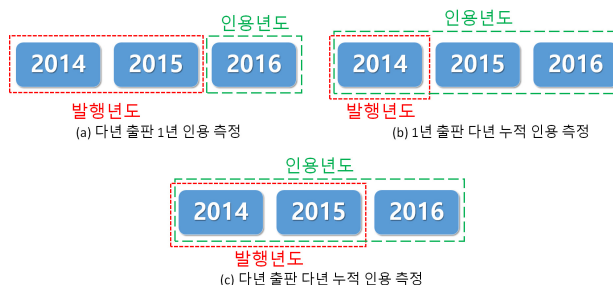
4.1 WoS 데이터와 인용 측정 방식

이 장에서는 완벽주의 지수 PI와 준완벽주의

지수 NPI를 학술지의 특성 파악에 실제로 적용해보기로 한다. 대상 데이터로 Web of Science(이하 WoS로 표기)의 'Information and Library Science' 분야 2017년 영향력 지수 상위 학술지 20종을 선택하였다. 이중에서 20위는 Journal of Enterprise Information Management이지만, 이 학술지가 2015년에 등재되어 그 이전 데이터가 없으므로 제외하고 21위인 Research Evaluation을 대상으로 포함하였다. 이들 20종 학술지에 대해서 WoS의 검색 기능으로 각 학술지에 2013년부터 2016년까지 게재된 논문을 모두 검색한 후, 각 논문이 해마다 인용된 횟수를 파악하였다. WoS에서는 검색 결과에서 'citation report' 기능을 활용하면 논문마다 발행된 이후 각 연도에 인용된 횟수를 파악할 수 있다.

PI와 NPI 산출에 있어서 논문의 출판년도 범위와 인용년도 범위를 어떻게 설정하는가에 따라서 <그림 5>와 같이 (a) 다년 출판 1년 인용 측정 방식, (b) 1년 출판 다년 누적 인용 측정 방식, (c) 다년 출판 다년 누적 인용 측정 방식이 가능하다.

다년 출판 1년 인용 측정 방식은 영향력 지수처럼 직전 n년 동안의 논문이 한 해 동안 인용된 횟수로 지수를 산출하는 방식으로서 영향력



<그림 5> 인용 지수 산출을 위한 논문의 발행년도 범위와 인용년도 범위 설정 방식

지수를 산출하는 방식과 같다. 예를 들어 <그림 5>의 (a)처럼 2014년과 2015년에 발행된 논문이 2016년에 인용된 횟수를 사용하여 평균 인용 횟수와 각 지수를 측정할 수 있다.

다년 누적 인용 측정 방식은 특정 연도 또는 일정 기간 동안 발행된 논문이 마지막 발행년도로부터 n 년 뒤(예를 들면 2년 뒤)까지 인용된 누적 횟수를 사용해서 지수를 산출하는 것이다. 이렇게 누적 인용횟수를 사용하면 1년 인용 측정 방식에 비해 논문이 인용될 시간적인 기회를 더 많이 얻게 된다. 따라서 많은 인용을 받는 논문도 발생하고 인용횟수 분포 곡선이 제대로 형성되어 지수 측정이 더 원활해질 수 있다. 다년 누적 인용 측정 방식은 출판년도를 한 해로 하는 1년 출판 다년 누적 인용 측정 방식과 여러 해 동안 발행된 논문을 대상으로 하는 다년 출판 다년 누적 인용 측정 방식으로 나눌 수 있다.

1년 출판 다년 누적 인용 측정 방식은 특정 연도 한 해 동안 발행된 논문이 발행년도로부터 n 년 뒤까지 인용된 횟수를 사용해서 지수를 산출하는 것이다. 예를 들어 <그림 5>의 (b)와 같이 2014년에 발행된 논문이 2년 뒤인 2016년까지 인용된 누적 횟수로 지수를 산출할 수 있다. 이는 발행년도가 한 해로 고정되고 인용년도가 여러 해에 걸치므로 통시적 영향력지수(Diachronous Impact Factor)(이재윤, 2018; Ingwersen, Larsen, Rousseau, & Russell, 2001)와 유사한 방식이다. 그런데 사전 실험 결과 한 해 동안에 발표된 논문만 사용하게 되면 논문 수가 적은 학술지의 경우에는 지수가 0에서 크게 벗어나지 않았다. Sidiropoulos, Katsaros, Manolopoulos(2015)가 완벽주의 지수 PI를 제

안했을 때에도 논문 수가 어느 정도 이상이 되어야 유력자나 대량생산자임을 파악할 수 있다고 보고하였다. 따라서 학술지에 대해 PI와 NPI를 측정할 때 대상 논문 수가 너무 적을 수도 있는 1년 출판 다년 누적 인용 측정 방식은 바람직하지 않다. 따라서 이 연구에서는 1년 출판 다년 누적 인용 측정 방식은 사용하지 않았다.

다년 출판 다년 누적 인용 측정 방식은 여러 해(주로 WoS의 영향력지수처럼 2년에서 5년 사이)에 걸쳐 출간된 논문이 평가시점 연도까지 인용된 누적 횟수를 사용해서 지수를 산출하는 방식이다. 예를 들면 <그림 5>의 (c)와 같이 논문의 발행년도를 2014년과 2015년의 2년 동안으로 지정하되 발행 후 2016년까지 누적 인용된 횟수로 지수를 산출한다. 평가 대상 논문의 발행년도를 한 해가 아닌 여러 해로 정한다는 점은 영향력 지수와 같지만, 인용시기를 한 해로 고정하는 영향력 지수와 달리 발행 후 평가 기준 연도까지 누적된 횟수를 사용하여 지수를 산출하는 점은 통시적 영향력지수와 같다. 이렇게 발행기간을 여러 해로 적용함으로써 인용횟수 분포를 측정하기에 필요한 최소한의 논문 수를 확보하고, 인용기간도 여러 해로 적용함으로써 각 논문의 가치가 후속 논문에 반영될 시간도 충분히 확보할 수 있다.

4.2 측정 방식에 따른 지수 산출 결과 비교

4.2.1 다년 출판 1년 인용 측정을 통한 지수 산출 결과

2년 단위 영향력 지수를 산출하는 방식과 마찬가지로 WoS 문헌정보학 분야 20종 학술지가 2014년과 2015년에 발행된 논문이 2016년

에 인용된 횟수를 사용하여 평균 인용횟수와 각 지수를 측정해보았는데, 결과는 <표 2>와 같다. 이 표에서 P는 논문 수, C는 총 인용횟수, C/P는 평균 인용횟수, h는 h-지수이다.

2년 동안 발행된 논문이 마지막 발행년도와 차년도 1년 동안만 인용된 횟수로 산출한 경우에는 <표 2>에서 보듯이 모든 학술지의 PI가 음수로 산출되었다. h-꼬리 부분에 부여하는 패널티를 차등적으로 부여하는 NPI에서는 전체의 1/4인 5종의 학술지만 양수로 산출되었

다. 또한 h-지수에 해당하는 h-코어 논문이 전체 논문에서 차지하는 비율이 최대 15.9%에서 최소 2.6%에 불과한 경우까지 있어서 h-꼬리에 해당하는 논문이 전체 논문 중 최대 97.4%를 차지하는 경우도 있는 것으로 나타났다. 이는 인용될 기회를 충분히 얻지 못한 논문이 매우 많기 때문에 나타나는 현상이다.

결국 1년 인용 측정 방식으로 한 해에만 인용되는 횟수를 사용하면 h-코어를 넘는 초과 인용이 발생할 기회가 적게 되고 h-꼬리에 속한 논

<표 2> WoS 문헌정보학 분야 학술지 20종의 2014년, 2015년 발행 논문이 2016년에 인용된 횟수로 산출한 지수

학술지	P	C	C/P	h	PI	NPI
INTERNATIONAL JOURNAL OF COMPUTER-SUPPORTED COLLABORATIVE LEARNING	41	128	3.12	6	-82	23.0
JOURNAL OF STRATEGIC INFORMATION SYSTEMS	44	142	3.23	7	-117	12.5
INFORMATION SYSTEMS JOURNAL	54	172	3.19	7	-157	7.5
JOURNAL OF INFORMATION TECHNOLOGY	62	150	2.42	6	-186	-18.0
JOURNAL OF THE ASSOCIATION FOR INFORMATION SYSTEMS	67	137	2.04	6	-229	-46.0
RESEARCH EVALUATION	67	161	2.40	7	-259	-49.0
EUROPEAN JOURNAL OF INFORMATION SYSTEMS	84	221	2.60	7	-318	-48.5
JOURNAL OF COMPUTER-MEDIATED COMMUNICATION	98	472	4.82	9	-329	71.5
SOCIAL SCIENCE COMPUTER REVIEW	102	191	1.87	6	-385	-97.0
INFORMATION PROCESSING & MANAGEMENT	120	293	2.44	7	-498	-102.5
JOURNAL OF KNOWLEDGE MANAGEMENT	134	349	2.60	7	-540	-95.5
JOURNAL OF MANAGEMENT INFORMATION SYSTEMS	105	219	2.09	8	-557	-169.0
MIS QUARTERLY	107	723	6.76	15	-657	33.0
TELEMATICS AND INFORMATICS	138	503	3.64	10	-777	-137.0
GOVERNMENT INFORMATION QUARTERLY	154	557	3.62	11	-1016	-229.5
INFORMATION & MANAGEMENT	168	671	3.99	11	-1056	-192.5
INTERNATIONAL JOURNAL OF INFORMATION MANAGEMENT	163	686	4.21	12	-1126	-220.0
JOURNAL OF INFORMETRICS	174	529	3.04	11	-1264	-367.5
JOURNAL OF THE ASSOCIATION FOR INFORMATION SCIENCE AND TECHNOLOGY	431	990	2.30	11	-3630	-1320.0
JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION	418	1403	3.36	13	-3862	-1229.5

* PI값 내림차순 정렬
 ** 지수별 1위 값은 음영 표시

문도 인용될 기회가 너무 적어서 패널티를 받게 될 가능성이 더 커지게 된다. 즉, 발행된 논문의 가치를 충분히 파악할 시간이 부족하여 인용횟수 분포 곡선이 제대로 형성되지 못하므로 PI나 NPI로 유력 학술지를 파악하기가 어렵다.

4.2.2 다년 누적 인용 측정을 통한 지수 산출 결과

- 1) 2년 출판 누적 인용 측정 결과
- 1년 인용 측정 방식에 비해서 인용될 기회를

더 부여하기 위해서 직전 2년 동안의 논문이 발행 직후부터 평가년도까지 인용된 누적 횟수를 사용하는 경우를 살펴보았다. 이럴 경우는 평가 대상 중 마지막 호에 발행된 논문은 인용될 기회가 다음 해 1년으로 1년 인용 측정 방식에서와 같지만, 평가 대상 중 첫 호에 발행된 논문은 당해년도부터 2년 뒤까지 거의 3년 가까이 인용될 기회를 가지게 된다. 2014년과 2015년에 발행된 논문이 2016년까지 인용된 누적 횟수를 사용하여 각 지수를 측정한 결과는 <표 3>과 같다.

<표 3> WoS 문헌정보학 분야 학술지 20종의 2014~2015년 발행 논문이 2016년까지 인용된 누적 횟수로 산출한 지수(PI 값 내림차순 정렬)

학술지	P	C	C/P	h	PI	NPI
INTERNATIONAL JOURNAL OF COMPUTER-SUPPORTED COLLABORATIVE LEARNING	41	176	4.29	7	-62	57.0
JOURNAL OF STRATEGIC INFORMATION SYSTEMS	44	227	5.16	9	-88	69.5
INFORMATION SYSTEMS JOURNAL	54	295	5.46	9	-110	92.5
JOURNAL OF INFORMATION TECHNOLOGY	62	208	3.35	8	-224	-8.0
JOURNAL OF THE ASSOCIATION FOR INFORMATION SYSTEMS	67	205	3.06	8	-267	-31.0
JOURNAL OF COMPUTER-MEDIATED COMMUNICATION	98	759	7.74	12	-273	243.0
RESEARCH EVALUATION	67	236	3.52	9	-286	-25.0
EUROPEAN JOURNAL OF INFORMATION SYSTEMS	84	377	4.5	9	-298	39.5
SOCIAL SCIENCE COMPUTER REVIEW	102	284	2.78	8	-468	-92.0
MIS QUARTERLY	107	1028	9.61	17	-502	263.0
INFORMATION PROCESSING & MANAGEMENT	120	445	3.71	9	-554	-54.5
JOURNAL OF KNOWLEDGE MANAGEMENT	134	461	3.44	9	-664	-101.5
JOURNAL OF MANAGEMENT INFORMATION SYSTEMS	105	348	3.31	12	-768	-210.0
TELEMATICS AND INFORMATICS	138	827	5.99	14	-909	-41.0
INTERNATIONAL JOURNAL OF INFORMATION MANAGEMENT	163	1058	6.49	14	-1028	15.0
GOVERNMENT INFORMATION QUARTERLY	154	765	4.97	13	-1068	-151.5
JOURNAL OF INFORMETRICS	174	809	4.65	13	-1284	-237.5
INFORMATION & MANAGEMENT	168	977	5.28	15	-1318	-170.5
JOURNAL OF THE ASSOCIATION FOR INFORMATION SCIENCE AND TECHNOLOGY	431	1697	3.94	16	-4943	-1623.0
JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION	418	2515	6.02	21	-5822	-1653.5

* PI값 내림차순 정렬

** 지수별 1위 값은 음영 표시

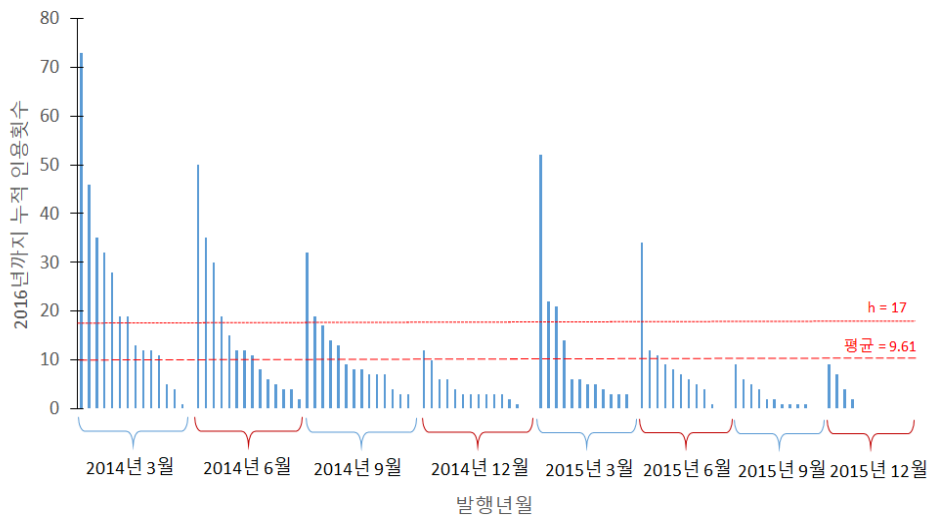
2년 동안 출판된 논문의 누적 인용횟수를 기준으로 지수를 산출하면 앞 연도에 출간된 논문은 상대적으로 인용될 시간적인 기회를 더 얻게 된다. 그러나 <표 3>에서 보듯이 1년 인용 측정 방식에서와 마찬가지로 PI는 여전히 모든 학술지가 음수로 나타나서 유력 학술지를 파악할 수가 없었다. 반면에 꼬리 영역 보완 패널티를 차등적으로 부여하는 NPI 기준으로는 263.0으로 가장 유력 학술지로 측정된 *MIS Quarterly*를 비롯해서 7종의 학술지가 양수로 산출되었다.

<표 3>에서 논문 수와 각 지수의 스피어맨 순위 상관계수를 측정해보면 PI는 -0.981로 나타나서 학술지의 PI 순위가 논문 수에 거의 완전하게 반비례하는 것으로 나타났다. PI 기준으로 1위와 2위를 차지한 학술지는 논문 수가 가장 적은 20위와 19위인 학술지였다. 한편 NPI와 논문 수의 스피어맨 순위 상관계수는 PI의 경우보다 다소 상관관계가 약한 -0.730이었고, NPI 기준 1위와 2위인 학술지는 논문 수로 중

간 정도인 10위(*MIS QUARTERLY*)와 13위(*JOURNAL OF COMPUTER-MEDIATED COMMUNICATION*) 학술지였다. 이를 통해 완벽주의 지수 PI는 논문 수에 지나치게 크게 좌우되지만, 준완벽주의 지수 NPI는 논문 수에 좌우되는 정도가 상대적으로 덜한 것이 드러났다.

2년에 가까운 발행시기의 차이가 각 논문의 인용횟수에 어떤 영향을 끼치는지 구체적으로 알아보기 위해서 평균 인용횟수가 9.61로 가장 높은 *MIS QUARTERLY*의 2014년, 2015년 두 해 동안 발간된 논문이 2016년까지 얼마나 인용되었는가를 <그림 6>에 제시하였다. 이 그림을 보면 발행시기의 끝 부분인 2015년 9월과 12월에 발간된 논문은 모두가 h-지수값 17에 미치지 못할 뿐만 아니라 평균인 9.61에도 미치지 못하여 PI 값을 산출할 때 상당한 패널티를 발생시키고 있다.

따라서 영향력 지수처럼 지수 산출 대상 논문의 출판년도를 직전 2년(또는 직전 n년)으로



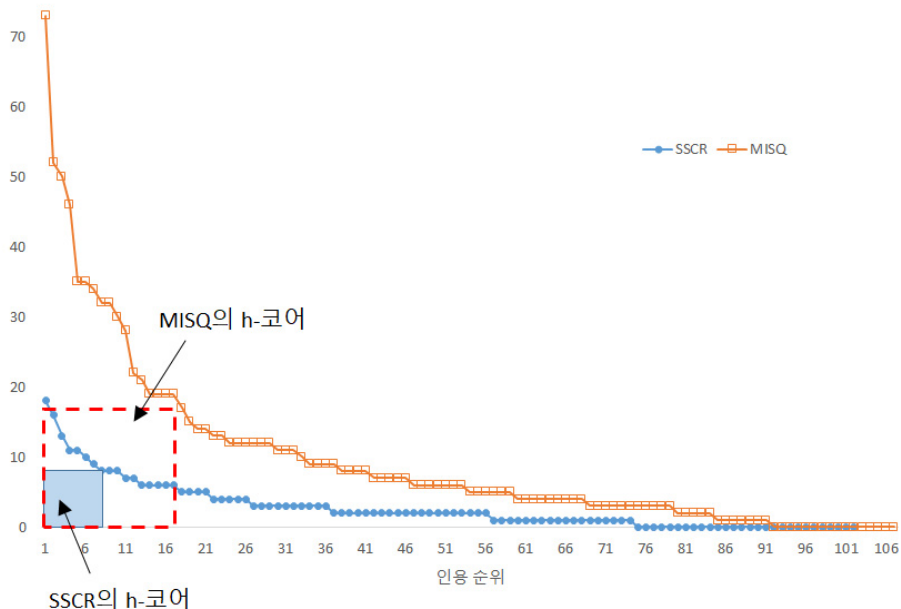
<그림 6> MIS Quarterly의 2014년, 2015년 발행 논문의 2016년까지 누적 인용횟수

설정할 경우에는 인용횟수는 출판 이후 평가년도까지의 누적횟수로 산출하고(2년 출판-누적 인용 방식), h-꼬리에 포함되는 논문에 부여하는 페널티는 NPI 방식으로 설정하는 것이 바람직하다.

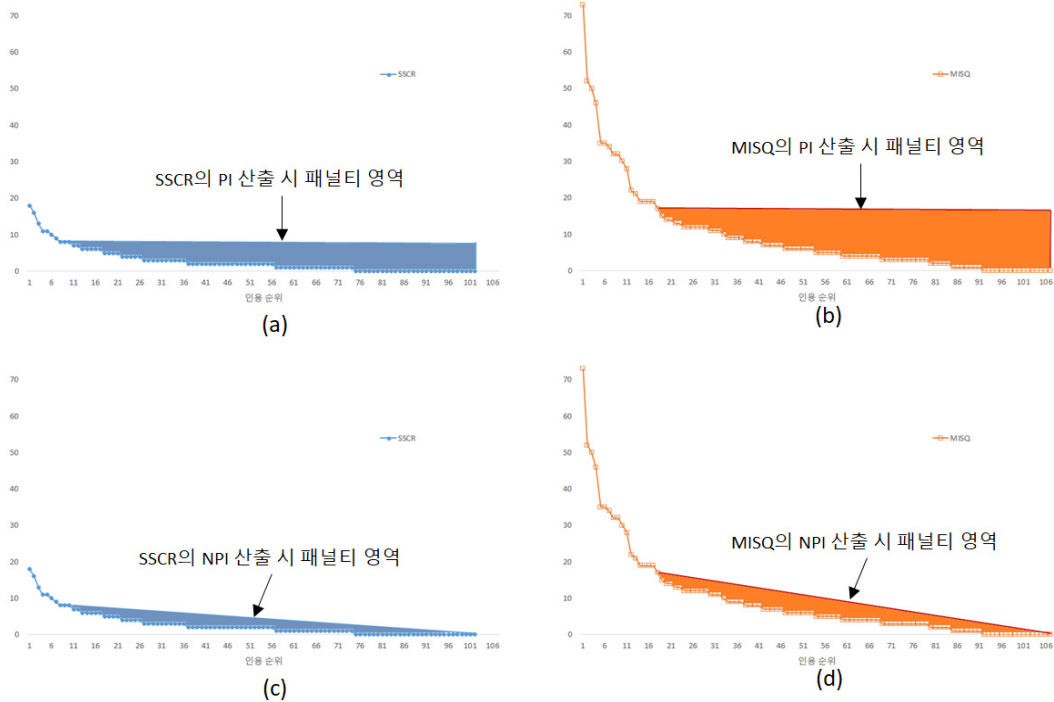
흥미로운 것은, 2년 동안의 발행 논문수가 각각 102편과 107편으로 비슷한 규모의 *JOURNAL OF MANAGEMENT INFORMATION SYSTEMS*(이하 JMIS)와 *MIS QUARTERLY*(이하 MISQ)의 순위가 PI 기준과 NPI 기준에서 뒤바뀌어 산출되었다는 점이다. 2년 동안 JMIS는 102편의 논문을 발행하여 2016년까지 평균 2.78회 인용되고 h-지수가 8인 반면, MISQ는 107편의 논문을 발행하여 2016년까지 평균 9.61회 인용되고 h-지수가 17이었다. 즉, 발행된 논문의 평균 인용 수준은 MISQ가 JMIS보다 3.5배 정도로 월등하게 높지만 PI 기준으로는 JMIS

가 더 높게 평가되었다. 반면에 NPI 기준으로는 MISQ가 더 높게 평가되었다. 이런 차이가 나타난 원인은 두 학술지에서 2년 동안 발행된 논문의 인용횟수 분포를 나타낸 <그림 7>을 보면 알 수 있다.

<그림 7>에서 보듯이 SSCR의 h-지수는 8이고 MISQ의 h-지수는 17이어서 h-코어의 크기는 MISQ(289)가 SSCR(64)에 비해서 4배 이상 더 크다. 그러나 이렇게 큰 h-지수 때문에 PI 산출 과정에서 MISQ의 페널티 영역은 1,068로 SSCR의 564보다 504나 더 큰 페널티를 받게 된다(<그림 8>의 (a)와 (b) 참조). 반면에 인용횟수의 선형적인 감소를 이상적인 목표 수준으로 하는 NPI 산출 과정에서는 MISQ의 페널티 영역이 303으로 SSCR의 페널티 영역 188에 비해서 큰 정도가 115에 불과하다(<그림 8>의 (c)와 (d) 참조). 따라서 순위가 낮을수록 인용횟



<그림 7> SSCR의 h-core(작은 사각형)와 MISQ의 h-코어(큰 사각형) 비교



〈그림 8〉 SSCR과 MISQ의 PI와 NPI 산출 시 패널티 영역 비교

수가 감소한다는 자연스러운 현상을 전제로 한 NPI 기준으로는 MISQ가 263, SSCR이 -92로 상호간의 우열이 뒤바뀌게 된다.

2) 3년 출판 누적 인용 측정 결과

발행 기간을 1년 더 확장하여 3년 동안 발행된 논문의 누적 인용횟수로 지수를 산출해보면 PI와 NPI의 차이를 더욱 잘 알 수 있다. 출판 범위를 3년으로 할 경우 2년인 경우에 비해서 전체 논문 수가 증가했고 상위 논문의 인용기회가 1년 더 늘어났음에도 불구하고 <표 4>를 보면 PI는 여전히 모두 음수로 산출될 뿐만 아니라 2년 동안 발행된 논문의 누적 인용횟수로 평가했을 때보다 값이 더 감소한 것으로 나타났다.

구체적으로 다년 출판 누적 인용 측정 방식

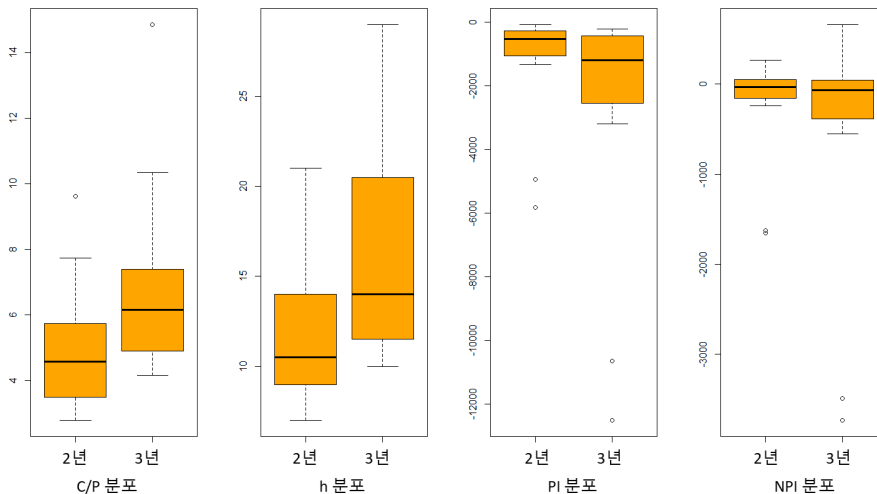
에서 출판 기간이 2년인 경우와 3년인 경우의 평균 인용횟수(C/P), h-지수, PI, NPI의 값 분포를 각각 비교해보면 <그림 9>와 같다. 3년치 논문의 누적 인용횟수로 지수를 산출하면 2년치 논문을 대상으로 한 경우보다 평균 인용횟수와 h-지수는 상승하는데, PI 값은 오히려 크게 감소한 것으로 나타난다. 결국 평가 대상 기간을 넓히면 PI값은 모든 학술지가 0 이하의 낮은 수준으로 더욱 하락하므로 학술지에 대해서 PI를 산출하여 유력 학술지와 대량생산 학술지를 구분하기는 어렵다. 한편 NPI 값은 두 결과의 중앙값(상자 가운데 굵은 가로 실선)은 모두 0에 가깝게 비슷한 수준을 유지하며, 2년치 논문을 대상으로 했을 때 보다 3년치 논문으로 측정했을 때 상위 학술지의 NPI는 더 상승

〈표 4〉 WoS 문헌정보학 분야 학술지 20종의 2013~2015년 발행 논문이 2016년까지 인용된 누적 횟수로 산출한 지수

학술지	P	C	C/P	h	PI	NPI
INTERNATIONAL JOURNAL OF COMPUTER-SUPPORTED COLLABORATIVE LEARNING	63	327	5.2	10	-203	62.0
JOURNAL OF STRATEGIC INFORMATION SYSTEMS	70	413	5.9	11	-236	88.5
INFORMATION SYSTEMS JOURNAL	86	560	6.5	12	-328	116.0
JOURNAL OF COMPUTER-MEDIATED COMMUNICATION	126	1303	10.3	15	-362	470.5
JOURNAL OF INFORMATION TECHNOLOGY	86	372	4.3	10	-388	-8.0
JOURNAL OF THE ASSOCIATION FOR INFORMATION SYSTEMS	96	468	4.9	11	-467	0.5
RESEARCH EVALUATION	99	411	4.2	10	-479	-34.0
EUROPEAN JOURNAL OF INFORMATION SYSTEMS	127	809	6.4	14	-773	18.0
JOURNAL OF MANAGEMENT INFORMATION SYSTEMS	150	903	6.0	14	-1001	-49.0
SOCIAL SCIENCE COMPUTER REVIEW	154	687	4.5	13	-1146	-229.5
TELEMATICS AND INFORMATICS	178	1057	5.9	14	-1239	-91.0
MIS QUARTERLY	174	2586	14.9	26	-1262	662.0
JOURNAL OF KNOWLEDGE MANAGEMENT	190	1196	6.3	15	-1429	-116.5
INFORMATION PROCESSING & MANAGEMENT	212	1026	4.8	14	-1746	-360.0
INFORMATION & MANAGEMENT	229	1659	7.2	19	-2331	-336.0
GOVERNMENT INFORMATION QUARTERLY	223	1682	7.5	22	-2740	-529.0
JOURNAL OF INFORMETRICS	277	1778	6.4	18	-2884	-553.0
INTERNATIONAL JOURNAL OF INFORMATION MANAGEMENT	266	2387	9.0	23	-3202	-407.5
JOURNAL OF THE ASSOCIATION FOR INFORMATION SCIENCE AND TECHNOLOGY	651	3187	4.9	22	-10651	-3732.0
JOURNAL OF THE AMERICAN MEDICAL INFORMATICS ASSOCIATION	652	5543	8.5	29	-12524	-3490.5

* PI값 내림차순 정렬

** 지수별 1위 값은 음영 표시



〈그림 9〉 다년 출판 누적 인용 측정 방식에서 출판 기간이 2년인 경우와 3년인 경우의 WoS 문헌정보학 분야 20종 학술지의 각 지수값 분포 비교

하고 하위 학술지의 NPI는 더욱 하락하여 학술지 간 변별력이 커지는 것을 보여준다. 따라서 PI와는 다르게 NPI는 평가 대상 연도 범위를 넓히더라도 높은 양의 값을 가지는 유력 학술지와 낮은 음의 값을 가지는 대량생산 학술지를 구분하는데 유용한 지수값을 산출할 수 있음이 확인된다.

5. 결 론

개인 연구자의 특성을 파악하기 위한 완벽주의 지수 PI와 이를 변형한 준완벽주의 지수 NPI로 학술지의 특성을 파악하는 시도를 해보았다. 결과적으로 PI는 발행 기간을 늘여서 논문 수를 늘이거나 인용 기간을 늘여서 인용 수준을 높여도 오히려 지수가 음수로 더 낮아져서 학술지의 특성을 파악하는데 유용하지 않은 것으로 나타났다. 반면에 이 연구에서 PI를 개선하여 제안한 NPI는 논문 수를 늘이거나 인용 수준을 높일수록 유력 학술지와 대량생산 학술지를 더욱 잘 구분하게 되는 것으로 나타났다. 3년 이상의 기간 동안 발행된 논문에 대해서 발행 기간 마지막 해의 차년도까지 인용된 누적 횟수로 NPI를 산출하면 적절하다고 판단된다. 예를 들어 2019년까지 누적 인용횟수로 지수를 산출한다면 2016년부터 2018년까지 발행된 논문이 인용된 횟수로 지수를 산출해야 할 것이다.

시험 적용 결과에서 완벽주의 지수 PI가 학술지 평가에 적합하지 않은 이유에는 연구자 평가와 학술지 평가의 특성이 다른 것도 한 원인으로 짐작된다. h-지수를 활용한 개인 연구자

평가에서는 연구자가 발표한 논문의 발행 시기를 특별히 제한하지 않고 오래전 발표한 논문이 최근까지 인용된 누적 횟수가 모두 지수에 반영되므로 최상위권 인용 논문이 가질 수 있는 인용횟수에 한계가 없다. 반면에 학술지 평가는 평가년도 기준 직전 2년에서 5년 정도 이내에 발행된 논문을 대상으로 하고 비교적 최근의 인용 위주로 평가하기 마련이므로, 기간 내 학술지 논문 중에서 가장 많이 인용된 최상위 논문이라 하더라도 어느 정도 이상으로 인용되기는 어렵다. 따라서 연구자 개인은 h-지수를 넘는 초과영역의 크기가 상당히 커질 수 있는데 반해서 학술지의 경우에는 h-지수에 비해 초과영역의 크기가 어느 정도 이상 커지기 어렵다.

어떤 연구자가 유력자가 되거나 대량생산자가 되게 하는 주된 원인은 연구자 자신의 논문 발표 행태이듯이, 어떤 학술지가 유력 학술지가 되거나 대량생산 학술지가 되게 하는 주된 원인은 학술지 편집진의 논문 심사 원칙이다. 학문적 수월성을 추구하면서 내용의 참신성, 방법의 타당성, 학문발전에 대한 기여도 등 엄격한 심사 기준을 유지한다면 게재 논문 수는 증가하지 않더라도 유력 학술지로 인정받을 수 있을 것이다.

이 연구에서 제시된 준완벽주의 지수 NPI는 학술지의 다양한 특성을 파악하는데 기여할 것으로 기대된다. 평가 기준이 비교적 단순한 국내 KCI 등재지에 대해서도 제안된 지수를 적용하여 다면적인 평가를 시도하는 것이 바람직하다. 또한 국제 학계에 비하면 국내 학계는 연구자의 수나 학술 논문의 수가 적으므로 인용횟수 분포에서 h-지수를 넘는 초과 영역이 크게 성장하기가 어렵기 때문에 PI가 아닌 NPI를 국내 연구자 평가에 적용해볼 필요도 있다.

참 고 문 헌

- 김관준, 이재운 (2010). 학술지 영향력 측정을 위한 h-지수의 응용에 관한 연구. 정보관리학회지, 27(1), 269-287. <http://doi.org/10.3743/KOSIM.2010.27.1.269>
- 민윤경, 차미경 (2017). 공공연구기금 논문성과물 게재 학술지의 오픈액세스 정책에 대한 연구. 정보관리학회지, 34(1), 155-176. <http://doi.org/10.3743/KOSIM.2017.34.1.155>
- 박지연, 김정은, 민윤경 (2010). 기관단위 연구성과 평가방법에 관한 연구. 정보관리학회지, 27(1), 249-267. <http://doi.org/10.3743/KOSIM.2010.27.1.249>
- 유소영, 이재운, 정은경, 이보람 (2015). 연구성과평가 지침 리뷰 및 국내 적용 제안을 위한 고찰. 정보관리학회지, 32(4), 249-272. <http://doi.org/10.3743/KOSIM.2015.32.4.249>
- 이재운 (2018). 출판년도의 즉시 인용빈도를 포함하는 학술지 인용지수 개발. 한국문헌정보학회지, 52(4), 71-90. <https://doi.org/10.4275/KSLIS.2018.52.4.071>
- Baum, J. A. C. (2011). Free-riding on power laws: Questioning the validity of the impact factor as a measure of research quality in organization studies. *Organization*, 18(4), 449-466. <https://doi.org/10.1177/1350508411403531>
- Björk, B. (2015). Have the “mega-journals” reached the limits to growth? *PeerJ*, 3, e981. <http://doi.org/10.7717/peerj.981>
- Braun, T., Glänzel, W., & Schubert, A. (2006). A hirsch-type index for journals. *Scientometrics*, 69(1), 169-173. <http://doi.org/10.1007/s11192-006-0147-4>
- Brzezinski, M. (2015). Power laws in citation distributions: Evidence from scopus. *Scientometrics*, 103(1), 213-228. <https://doi.org/10.1007/s11192-014-1524-z>
- Cole, S., & Cole, J. R. (1967). Scientific output and recognition: Study in operation of reward system in science. *American Sociological Review*, 32(3), 377-390. <http://doi.org/10.2307/2091085>
- Costas, R., & Bordons, M. (2008). Is g-index better than h-index? An exploratory study at the individual level. *Scientometrics*, 77(2), 267-288. <http://doi.org/10.1007/s11192-007-1997-0>
- Donner, P. (2018). Effect of publication month on citation impact. *Journal of Informetrics*, 12(1), 330-343. <https://doi.org/10.1016/j.joi.2018.01.012>
- Harzing, A.-W., & van der Wal, R. (2009). A google scholar h-index for journals: An alternative metric to measure journal impact in economics and business. *Journal of the American Society for Information Science and Technology*, 60(1), 41-46. <http://doi.org/10.1002/asi.20953>
- Hirsch, J. E. (2005). An index to quantify an individual's scientific research output. *Proceedings*

- of the National Academy of Sciences, 102(46), 16569-16572.
<http://doi.org/10.1073/pnas.0507655102>
- Ingwersen, P., Larsen, B., Rousseau, R., & Russell, J. (2001). The publication-citation matrix and its derived quantities. *Chinese Science Bulletin*, 46(6), 524-528.
<https://doi.org/10.1007/bf03187274>
- PLOS ONE (n.d.) PLOS ONE journal information. Retrieved from
<https://journals.plos.org/plosone/s/journal-information>
- Redner, S. (1998). How popular is your paper? An empirical study of the citation distribution. *The European Physical Journal B*, 4(2), 131-134. <http://doi.org/10.1007/s100510050359>
- Rosenberg, M. S. (2014). A biologist's guide to impact factors. *PeerJ Preprints* 2:e477v1. Retrieved from <https://peerj.com/preprints/477/>
- Sidiropoulos, A., Katsaros, D., & Manolopoulos, Y. (2015). Ranking and identifying influential scientists versus mass producers by the perfectionism index. *Scientometrics*, 103(1), 1-31.
<http://doi.org/10.1007/s11192-014-1515-0>
- Spezi, V., Wakeling, S., Pinfield, S., Creaser, C., Fry, J., & Willett, P. (2017). Open-access mega-journals: The future of scholarly communication or academic dumping ground? A review. *Journal of Documentation*, 73(2), 263-283. <http://doi.org/10.1108/JD-06-2016-0082>
- van Raan, A.F.J. (2004). Sleeping beauties in science. *Scientometrics*, 59(3), 467-472.
<https://doi.org/10.1023/B:SCIE.0000018543.82441.f1>
- Ye, F. Y., & Rousseau, R. (2010). Probing the h-core: An investigation of the tail-core ratio for rank distributions. *Scientometrics*, 84(2), 431-439.
<http://doi.org/10.1007/s11192-009-0099-6>

• 국문 참고문헌에 대한 영문 표기
(English translation of references written in Korean)

- Kim, Pan-Jun, & Lee, Jae Yun (2010). A study on journal impact measurement with hirsch-type indices. *Journal of the Korean Society for Information Management*, 27(1), 269-287.
<http://doi.org/10.3743/KOSIM.2010.27.1.269>
- Lee, Jae Yun (2018). Developing new journal citation indicators including immediate citation frequencies in the published year. *Journal of the Korean Society for Library and Information Science*, 52(4), 71-90. <https://doi.org/10.4275/KSLIS.2018.52.4.071>
- Min, Yoonkyung, & Cha, Mikyeong (2017). A study on the open access policy of scholarly

journals publishing research papers funded by Korean government. *Journal of the Korean Society for Information Management*, 34(1), 155-176.

<http://doi.org/10.3743/KOSIM.2017.34.1.155>

Park, Ji-Yeon, Kim, Jung-Eun, & Min, Yoon-Kyung (2010). A study on the evaluation methods of research institution: Based on the h-index and its variants. *Journal of the Korean Society for Information Management*, 27(1), 249-267. <http://doi.org/10.3743/KOSIM.2010.27.1.249>

Yu, So-Young, Lee, Jae Yun, Chung, EunKyung, & Lee, Boram (2015). A review of declarations on appropriate research evaluation for exploring their applications to research evaluation system of Korea. *Journal of the Korean Society for Information Management*, 32(4), 249-272. <http://doi.org/10.3743/KOSIM.2015.32.4.249>