

# 학술대회 논문의 참고자료 인용패턴 분석

- 정보과학 분야를 중심으로 -

## A Bibliometric Analysis of Citation Patterns in Conference Papers of Information Science

이다니엘 (Danielle Lee)\*

### 목 차

- |                          |                     |
|--------------------------|---------------------|
| 1. 서 론                   | 4. 표본논문 수집의 절차 및 출처 |
| 2. 참고자료의 인용패턴을 분석한 기존 연구 | 5. 분석 결과            |
| 3. 다양한 참고자료 관련 특성        | 6. 결 론              |

### 초 록

본 논문은 정보과학 분야를 중심으로 학술대회 논문에서 참고자료를 인용하는 패턴을 분석하고 피인용된 참고자료의 다양한 특성이 추후 그 자료를 참고한 학술대회 논문의 인용수에 어떤 영향을 끼치는지 분석하였다. 참고자료의 인용패턴에 관한 계량서지학 연구는 학술지 논문을 중심으로 활발히 이루어졌다. 하지만 점점 커지는 중요성에도 학술대회 논문을 대상으로 한 계량서지학 연구는 그 자체가 아직 초기 단계이다. 특히, 학술대회 논문에서 참고자료를 인용한 패턴이나 학술대회 논문의 인용수와 관련이 깊은 피인용 자료의 특성을 조사하는 연구는 미비하다. 따라서 본 연구는 1,900여편의 정보과학 관련 학술대회 논문을 표본논문으로 삼고 그 표본 논문에서 참고자료가 어떻게 인용되었는지 패턴을 살펴보았다. 또, 참고자료의 다양한 특성들(참고자료의 유형 및 개수, 피인용 당시 참고자료의 인용수, 피인용 당시 참고자료의 나이, 저자의 자기인용 비율)이 표본논문과 추후 얻게 될 인용수와 어떤 연관이 있는지 조사하였다. 그 결과, 피인용 참고자료의 53%가 학술대회 논문이었고, 29%가 학술지 논문이었다. 특히, 참고자료의 14%가 학술지 논문이나 단행본이 아닌 비전형적인 참고자료의 유형이었다. 그리고 참고자료의 60% 이상이 5년 이내에 발표된 자료이고, 최신자료일수록 학술대회 논문, 웹 페이지, 기타자료의 비중이 높았다. 참고자료 중 자기 인용 비율은 1.7%로 미비했다. 마지막으로 표본논문의 인용수에 유의미한 영향을 끼치는 참고자료의 특성은 참고자료로 쓰인 학술대회 논문의 수, 기타 자료의 수, 그리고 참고자료의 평균나이였다. 즉, 학술대회 논문과 기타자료를 참고자료로 많이 사용하고 최신의 자료를 많이 참조한 표본논문일수록 많이 인용되었다.

### ABSTRACT

This paper aims to explore the citation patterns of conference papers in 'Information Science' discipline and to analyze impacts of various cited works-related factors on future citations of conference papers. Existing bibliometric studies have investigated citation patterns and the statistical relations between a variety of bibliographic factors and the future citations of literature. However, the attentions have been focused largely on journal articles, and the bibliometric studies targeting conference papers are still in an infant stage. Therefore, this study, which is based on 1,904 conference papers in 'Information Science' field, examined several citation patterns and the contributions of the factors about cited works - the number of cited works, type of cited works, citation rates and ages of cited works at the time of being cited and the rate of self-citedness - to the future citation of the citing target articles. The data source of this study including the properties of target articles and cited works and citation rates of target articles was Scopus. As the results, 53% and 29% of the cited works were conference papers and journal articles, respectively. 14% of them are non-traditional types such as web pages, technical reports, patent, etc. More than 60% of the cited works were 5 years old or less. Among several factors considered in this paper, the number of conference papers and the number of non-traditional types of works are the most contributing factors on the citation rates of target articles. The recency of the cited works is also significant contributor on the citation rates of target articles. That is, the target articles citing more conference papers and non-traditional types of works earned more citations. The target articles citing recent works also earned more citations.

키워드: 참고자료, 인용패턴, 인용지수 분석, 계량서지학, 학술대회

Citation Analysis, References, Bibliometric Analysis, Multiple Regressions, Conference Proceedings

\* 상명대학교 공과대학 소프트웨어학과 조교수(suleehs@gmail.com)

논문접수일자: 2017년 10월 16일 최초심사일자: 2017년 10월 23일 게재확정일자: 2017년 11월 13일  
한국문헌정보학회지, 51(4): 35-52, 2017. [<http://dx.doi.org/10.4275/KSLIS.2017.51.4.035>]

## 1. 서 론

### 1.1 연구의 필요성

대다수 학문 분야에서는 연구결과물로서 학술지 논문(journal articles)을 중시하고 학술대회(conference papers)<sup>1)</sup> 논문을 등한시하는 반면(Nederhof 2006; Onodera and Yoshikane 2015; Scopus 2013; Tahamtan, Afshar and Ahamdzdeh 2016), 컴퓨터 과학에서는 학술대회 논문이 학술지 논문과 마찬가지로 중요하다. 여러 계량서지학 연구(Franceschet 2010; Freyne et al. 2010; Eckmann, Rocha and Wainer 2012; Vrettas and Sanderson 2015)가 컴퓨터 과학에서 학술대회 논문의 중요성을 학술지 논문과 비교하여 증명하였다. Freyne(Freyne et al. 2010)에 따르면 중요 학술대회 논문의 인용수가 중급 학술지 논문의 인용수와 비슷하였고, 하급 학술지 논문보다는 많았다. Vrettas (2015)에 따르면 학술대회 평가 체계 중 하나인 CORE ERA 순위에서 최상위인 A\*로 평가된 학술대회 논문은 같은 A\*로 평가된 학술지의 논문보다 더 많이 인용되었다(Vrettas and Sanderson 2015). 그리고 Wainer의 연구에 의하면, 2006년 한 해 동안 ACM 데이터베이스에 색인된 ‘컴퓨터 과학’ 관련 논문 중, 참고자료의 39%가 학술대회 논문이었고 10회 이상 피인용된 논문의 41%가 학술대회 논문이었다 (Wainer, Przibiszczki de Oliveira and Anido 2011).

컴퓨터 과학에서 이처럼 학술대회 논문이 중

요한 이유 중 하나는 빠르게 진화하는 학문 분야의 특성 때문이다. 논문 심사부터 발표까지의 기간이 학술지보다는 학술대회가 훨씬 짧다보니 중요한 연구결과를 학술대회에서 빨리 발표할 수 있다(Freyne et al. 2010; Vardi 2010; Eckmann, Rocha and Wainer 2012). 그리고 학술지 논문은 엄격한 심사절차를 통해 더욱 정교하고 견고한 논문 설계 및 실험을 요구하는 반면, 학술대회에서는 다양한 발표형태(구두발표, 포스터, 토론 등)로 인해 덜 정교하고 덜 성숙했지만 참신한 연구 아이디어를 발표할 수 있는 기회가 열려있다. 또한 동료 연구자들이 학술대회 기간 중에 서로의 연구에 대해 의견을 나눌 수 있고 대면활동으로 연구인맥을 넓혀 새로운 연구 기회를 모색할 수 있다는 이점도 있다(Freyne et al. 2010). 그리고 한 연구 주제의 논문을 학술지에 발표하기 이전에 간략하게 요약해서 먼저 학술대회에 발표하는 경우도 흔하다(Montesi and Owen 2008; Zhang and Jia 2013).

컴퓨터 과학과 정보과학 간의 활발한 상호학제간 연구교류로 인해 두 학문간의 학제 경계가 열어지고, 두 분야 모두 연구가 빠르게 진화한다는 점에서(Larivière, Sugimoto and Cronin 2012) 정보과학자들도 컴퓨터 과학자들과 마찬가지로 학술대회 논문을 중시한다. 그 증거로서 정보과학 분야의 서지정보를 연대기적으로 분석한 Larivière(Larivière, Sugimoto and Cronin 2012)는 21세기 들어 문헌정보학 논문의 참고자료 중 학술대회 논문의 비중이 5~6% 증가하였다고 보고했다. 이는 비교 대상이었던

1) 본 논문에서 ‘학술대회’는 논문 서술과 구두 발표를 영어로 하는 국제학술대회를 뜻하며 ‘학술대회 논문’은 국제 학술대회에서 발표된 논문을 뜻한다.

사회과학과 인류학의 증가치보다 두 배나 많다 (Larivière, Sugimoto and Cronin 2012).

이렇듯 연구결과물로서 컴퓨터과학과 정보 과학 분야의 학술대회 논문이 점점 중요해지지만, 학술대회 논문을 대상으로 삼는 계량서지학 연구는 국내외를 망라하고 아직 결음마단계이다. 학술지 논문 중심의 계량 서지학 연구는 다방면에 걸쳐 진행되고 있지만 학술대회에 대한 계량서지학 연구는 대체로 두 가지 측면에서 진행되었다. 우선은 학술지 논문과 비교해 학술 대회 논문의 질을 비교 평가하는 논문들로 구체적으로는 인용 수의 차이(Freyne et al. 2010; Vrettas and Sanderson 2015), 학술대회 논문을 학술지 논문으로 재생산되는 패턴(Montesi and Owen 2008; Wainer and Valle 2013), 참고자료의 양식 차이 등(Wainer, Przibiszcki de Oliveira and Anido 2011)을 연구하였다. 이는 학술대회 논문의 중요성을 선험적으로 증명하려는 노력이다. 또한 학술지 논문의 보편적 평가 기준인 JCR impact factors에 비견될 만한 국제학술대회의 평가기준이 없기 때문에 그 기준을 정하기 위한 연구들이기도 하다. 두번째 측면은 학술대회 논문의 여러 요소를 연대기적 측면에서 살펴보고 분석, 진단하는 연구이다. 이는 특정 연구 주제별, - 예를 들어, 데이터베이스 (Sakr and Alomari 2012), 국제관계학(Wuehrer and Smejkal 2013), human computer interaction (Bartneck and Hu 2009), 소프트웨어 공학 (Vasilescu et al. 2014) 등 - 특정 국가별, - 예를 들어, 브라질(Barbosa, Silveira and Gasparini 2017) - 혹은 특정 색인 데이터베이스 - 예를

들어, IEEE(Shirakawa et al. 2012) - 별로 관련 학술대회들이 발전한 모습을 분석하였다. 하지만 학술대회 논문 관련 세부 요소를 분석하고 나아가 그 요소가 학술대회 논문의 인용수와 어떤 관련이 있는지는 거의 연구하지 않았다 (Shirakawa et al. 2012; Tahamtan, Afshar and Ahamdzdeh 2016; Waltman 2016). 따라서 본 연구는 학술대회 논문 관련 요소 중 참고자료의 인용패턴을 분석하고 나아가 그 패턴들이 자료를 참고한 표본논문의 추후 인용수에 미치는 영향을 분석하였다.

## 1.2 연구의 목적 및 내용

본 연구는 다음의 질문들에 답하고자 한다. 학술지, 단행본, 학술대회 논문 중 학술대회 논문에서 참고자료<sup>2)</sup>로 많이 인용된 형태는 무엇인가? 매뉴얼, 특허, 웹페이지, 기술보고서 등 비전형적인 형태의 참고자료는 얼마나 인용되었고 비전형적인 참고자료가 많은 논문은 전형적인 참고자료가 많은 논문에 비해 특별히 영향력이 큰가? 최신의 아이디어를 발표할 수 있는 장점이 있는 학술대회 논문에서 최신의 참고자료를 사용하면 앞으로 받게 될 인용수에서 유리한가? 논문의 인용수를 높이기 위해 이미 많이 인용되어 명성을 쌓은 자료를 인용하는 것이 유리한가? 학술대회 논문에서 자신들의 논문을 참고자료로 인용하는 비율은 얼마나 되나? 자기 인용비율이 낮은 논문은 자기 인용비율이 높은 논문과 비교하여 얻게 될 인용수에 차이가 있는가? 이 질문들에 답하기 위해 1,904편의

2) 본 연구의 분석에 따르면 단행본, 학술지, 학술대회 논문 같은 문헌뿐만 아니라 웹 페이지, 특허, 뉴스레터, 기술보고서, 신문 기사, 매뉴얼 등 다양한 형태의 자료가 참고 되었다. 따라서 '참고문헌'보다는 '참고자료'로 서술하였다.

정보과학 분야 학술대회 논문을 표본논문으로 조사하였고, Scopus의 인용지수를 기준으로 분석을 실행하였다. 본 연구는 다음의 두 가지 측면에서 계량서지학 연구에 이바지한다. 첫 번째, 컴퓨터과학 및 정보과학 분야의 계량서지학 연구 중 특히 참고자료의 인용패턴을 분석하는 연구들이 대부분 학술지 논문에 편중되어 있지만, 본 연구는 그 연구대상을 학술대회 논문으로 확장한다. 두 번째, 학술대회 논문의 인용 수를 결정하는 요소가 무엇인지 분석하는 초기연구 중 하나이다.

다음 장에서는 기준에 검토된 참고자료 패턴들은 무엇인지 알아보고자 한다. 3장에서는 본 논문에서 검토한 다양한 참고자료의 특성들을 소개하고 4장에서는 표본논문 및 참고자료 수집 방법을 설명한다. 5장에서 분석 결과를 설명한 다음 마지막 6장에서 결론을 내릴 것이다.

## 2. 참고자료의 인용패턴을 분석한 기준 연구

국내 논문의 참고자료 인용패턴을 분석한 연구로는 우선 오세훈(2005)이 있다. 60년간 우리나라에서 발표된 문헌정보학 학술지 논문에서 참고자료의 인용패턴을 분석한 결과에 따르면 1990년대 초반을 기점으로 참고자료에서 해외 논문이 차지하는 비율이 현저히 증가하였다. 그리고 최근에 발표된 논문일수록 국내 논문보다는 해외 논문의 의존도가 높았다. 이는 연구자들의 외국어 능력이 높아지고 외국 자료를 접할 기회가 많아졌기 때문이라고 판단하였다. 특히 여러 세부분야 중에서도 정보과학이 자료

구조, 도서관 운영 등 다른 세부분야보다 2배정도 많이 해외논문을 참고자료로 인용하였다(오세훈 2005).

조인숙(2007)은 10년간 문헌정보학 관련 국내 학술대회지들에 실린 약 1,300편의 표본논문에서 인용된 참고자료의 형태를 분석하였다. 그 결과 각 논문당 평균 27편의 참고자료를 인용하였고 단행본과 학술지 논문의 비중이 각각 29%와 28%였다. 그리고 보고서(13%)와 웹페이지 순(10%)으로 나타났다. 특히 2000년 이후로 웹 페이지의 비중이 증가하였다. 이 연구에서 국내외 학술대회 논문은 참고자료의 형태 분석에서 제외하였는지, 아니면 참고자료 목록에서 전혀 발견되지 않았는지는 알 수 없다. 김미라(2013)는 국방과학 분야에서 약 15년간 발표된 약 500여편의 학술지 논문과 피인용 참고자료 4,500여편을 분석하였다. 그 참고자료 중 가장 많이 인용된 논문은 학술지(34%)였고, 연구보고서 및 기관간행물(26%), 단행본(20%), 그리고 학술대회 논문(16%)이 그 뒤를 이었다. 그리고 규격/표준과 특허 같은 비전형적인 참고자료도 약 4%였다. 비공개가 많은 국방과학분야의 특성으로 기존 학문과 참고자료 인용 패턴에 큰 차이를 발견하였다.

김갑선(2002)는 국내 사회과학자들이 참고자료를 선택하는 패턴을 정량적으로 분석했을 뿐 아니라 참고자료의 선택 동기를 설문조사를 통해 정성적으로도 분석하였다. 우선 설문조사에 참여한 학자들의 논문에 피인용된 참고자료 중 해외논문이 국내논문보다 2배 많았다. 그리고 피인용된 참고자료의 45%와 43%가 각각 단행본과 학술지였으며, 10%는 신문, 잡지, 심포지움 자료 등 정기간행물이었으며 웹 페이지

는 0.8%였다. 그리고 참고자료의 38%가 최근 5년내 발간된 최신논문이었고 자기인용 비율은 2.7로 미비했다. 설문 분석 결과에 따르면 연구 아이디어를 얻거나 개념적인 배경자료가 필요 할 때 주로 해외자료에 의존하고 이는 해당 연구자의 해외경험(학위 및 국외교육경험)이 많고 연구주제가 해외에서 활발히 진행되는 유형 일수록 의존도가 높았다.

이로써 국내 계량서지학 연구 중 참고자료의 인용 패턴을 분석한 연구들은 대부분 국내학술지 논문을 대상으로 하고, 참고자료로는 학술지논문과 단행본이 많이 인용되었고 해외 자료의 의존도가 점점 높아지는 추세임을 알 수 있

다. 하지만 본 연구의 표본논문은 학술대회 논문이고 참고자료도 모두 해외자료이므로 국내 계량서지학 연구의 결과와 접점을 찾기가 힘들었다. 따라서 해외 계량서지학 연구 중 참고자료의 인용패턴을 연구한 논문을 살펴보았다.

〈표 1〉은 피인용된 참고자료의 특성이 인용한 표본논문의 인용 수와 얼마나 관련 있는지를 요약하였다. 〈표 1〉에서 요약한 표본논문의 종류 및 분야에서 볼 수 있듯이, 거의 모든 연구의 대상이 학술지이다. 참고자료의 수(Bornmann et al. 2014; Chakraborty et al. 2014; Didegah and Thelwall 2013; Falagas et al. 2013; Onodera and Yoshikane 2015; van Wesel,

〈표 1〉 참고자료의 특성이 표본논문의 인용 수에 끼치는 영향력을 조사한 연구들

| 논문                                   | 논문형태(대상분야)                           | 표본 크기     | 조사한 저자특성   |
|--------------------------------------|--------------------------------------|-----------|--|
| Antoniou et al. (2015)               | 학술지<br>(혈관외과)                        | 226       | 주제, 연구설계, 제목의 유형, 논문의 길이, 참고자료의 유형, 저자 특성(지역적 특성 및 학제간 협력 정도), 논문의 오픈액세스 여부, 연구과제 여부   |
| Biscaro and Giupponi (2014)          | 학술지<br>(환경과학)                        | 3,343     | 공저자 네트워크에서의 중심성 척도(degree, betweenness, and closeness centrality), 참고자료 네트워크에서 서지결합(bibliographic coupling) 척도(degree, betweenness centrality, clustering coefficient) |
| Bornmann et al. (2014)               | 학술지<br>(WoS에 색인된 모든 학문분야)            | 475,391   | 출간이후 단기간(1년 혹은 2년)내에 쌓인 인용 수, 저자 수, 참고자료의 수, 논문의 길이  |
| Chakraborty et al. (2014)            | 학술지 및 학술대회<br>(컴퓨터 공학)               | 1,549,317 | 저자의 생산성, 저자의 H-index, 저자 연구의 학제간 다양성, 저자의 사회성, 출판지의 장/단기 명성, 출판지의 주제 다양성, 참고자료의 수, 참고자료의 학제간 다양성, 논문 주제의 다양성   |
| Didegah and Thelwall (2013)          | 학술지<br>(나노과학)                        | 50,162    | 학술지의 Impact factor, 저자 수, 저자의 국적, 참고자료의 인용 수, 참고자료의 수, 출간된 저널의 국제성(internationality) 정도, 참고자료의 국제성 정도  |
| Falagas et al. (2013)                | 학술지<br>(일반내과)                        | 196       | 저자수, 저자의 소속 기관 수, 제목과 요약의 단어 수, 논문의 길이, 참고자료 수, 연구 설계, 학술지의 Impact Factor  |
| Onodera and Yoshikane (2015)         | 학술지<br>(화학, 전기공학, 분자생물학, 물리학 & 소화기학) | 1,395     | 참고자료의 수, 참고자료의 Price Index, 저자 수, 소속기관 국적, 2000년 까지 주저자의 논문 수, 2000년까지 모든 저자 논문의 인용 수 & 주저자의 출판연력, 논문의 길이, 논문내의 그림, 표 & 수학수식의 수  |
| van Wesel, Wyatt and ten Haaf (2014) | 학술지<br>(사회학, 일반내과, 응용물리학)            | 32,649    | 제목의 길이, 요약의 길이, 논문의 길이, 참고자료의 수  |

(가장 오른쪽 행에서 밑줄이 있는 특성은 유의미한 영향력을 보였음)

Wyatt and ten Haaf 2014), 참고자료의 유형 (Antoniou et al. 2015), 참고자료의 인용네트워크에서 서지 결합의 사회적 척도(Biscaro and Giupponi 2014), 참고자료 내용의 학제간 혹은 저자 출신 국가의 다양성(Chakraborty et al. 2014; Didegah and Thelwall 2013), 참고자료의 최신성(Onodera and Yoshikane 2015), 참고자료의 인용수(Didegah and Thelwall 2013) 등을 연구하였다. 그러나 표에서도 알 수 있듯이, 학술지 논문들에서 조차 여러 학문 분야에 걸쳐 공동으로 유의미한 참고자료 특성은 없었다. 심지어 같은 학문분야에서도 공동으로 유의미한 특성은 없었다. 예를 들어, 피인용된 참고자료의 수가 표본논문의 인용수에 유의미한 영향을 끼친다는 주장(Onodera and Yoshikane 2015; van Wesel, Wyatt and ten Haaf 2014)도 있지만, 같은 대상 학문분야 내에서 조차 반대의 주장을 펴는 연구(Falagas et al. 2013)도 있었다.

Onodera(2015)는 Web of Science(WoS)에 색인된 학술지 중 여섯 가지 학문 분야를 선택하고 각 학문분야에서 저자 및 참고자료에 관한 다양한 특성을 분석하였다. 특히 그 다양한 특성들이 표본논문의 인용수에 어떤 영향을 끼치는지 살펴보았다. 그 결과 표본논문의 인용수를 가장 크게 좌우하는 것은 참고자료가 얼마나 최신의 자료인지를 나타내는 Price Index였다. 즉, 5년 이내에 출간된 자료를 많이 인용할수록 그 자료들을 참조한 표본논문도 많이 인용되었다. 또한 참고자료의 수도 유의미한 영향을 끼쳤다. 참고자료의 수가 많을수록 표본논문도 많이 인용되었지만, 논문의 길이는 인용수에 큰 영향이 없었다. 통상적으로 길이가 긴 논문

이 참고자료의 수도 많다고 생각하지만 이 연구에서는 논문의 길이와 참고자료의 개수간에 상관관계를 발견하지 못했다. Chakraborty(2014)는 컴퓨터 과학 분야의 학술대회 논문을 대상으로 한 연구 중 하나이다. 이 연구는 Microsoft Academic Search에서 찾은 1.5백만여 컴퓨터 과학분야의 논문을 분석하였다. 특히 참고자료의 수와 참고자료를 인용한 논문이 얼마나 다양한 주제로 쓰였는지 등 10가지가 넘는 논문 관련 특성을 분석하였다. 그리고 그 특성 중 어떤 것이 논문의 인용수를 가장 잘 예측할 수 있는지 기계학습을 통해 알아보았다. 하지만 저자 관련 특성이 인용수를 좌우하는 특성이었고, 참고자료 관련 특성은 유의미한 예측변수가 아니었다.

### 3. 다양한 참고자료 관련 특성

학술대회 논문에서 참고자료가 인용되는 패턴을 알아보려는 본 연구는 다음의 특성을 분석하였다.

- 각 표본논문 당 참고자료 수
- 참고자료의 유형
- 피인용 당시 참고자료의 나이
- 피인용 당시 참고자료의 인용횟수
- 참고자료 중 저자의 자기 인용횟수

위의 특성 중 학술지, 학술대회 논문, 단행본과 같이 전통적인 형태의 참고자료가 아닌 특허, 웹페이지, 기술보고서, 각종 표준의 규격, 매뉴얼 등 비전통적인 형태의 자료가 참고자료 목

록에서 차지하는 비율을 살펴본다. 그리고 그 개수가 추후 해당 표본논문이 얻게 될 인용수에 어떻게 영향을 끼치는가 분석한다. 참고자료의 나이는 참고자료의 발표년도부터 인용된 표본논문의 출간년도까지의 차이를 계산하였다. 예를 들어, 2006년에 발표된 논문이 2009년 CIKM 학술대회에서 발표된 한 논문에 인용되었다면 그 참고자료의 나이는 3년이다. Scopus에서는 출간 년도만 제공될 뿐 보다 상세한 월일 정보가 없기 때문에 년도기준으로 나이를 계산하였다. 이 특성을 통해 첨단의 연구 아이디어 및 결과를 발표할 수 있는 학술대회 논문에서 참고자료도 최신의 자료를 인용하는 것이 과연 추후 인용수를 높이는데 유리한가 살펴본다. 참고자료의 인용횟수는 피인용될 당시의 횟수를 계산하기 위해 Scopus의 API (<http://api.elsevier.com/>)를 통해 각 참고자료를 인용한 논문목록을 별도로 수집하고 그 목록 중 표본논문보다 먼저 출간된 논문의 수를 셨다. 앞서 예를 든 2009년 CIKM 논문에 참고자료로 인용된 2006년 논문은 2006년부터 2009년까지 피인용된 횟수를 계산하였다. 참고자료로 쓰일 당시의 인용수로 인해, 이미 많이 인용되어 영향력이 커진 논문을 참고자료로 인용할 때 과연 표본논문의 인용수에도 도움이 되는지 알아본다. 마지막으로 저자의 자기인용 비율은 표본논문 저자들의 출판 목록을 추가로 수집하여 Scopus에서 제공하는 저자 및 문서식별자를 이용해 자기인용을 계산하였다. 그리고 인용된 참고자료가 주저자의 논문이든, 공저자의 논문이든 모두 자기인용이라고 계산하

였다(Donner 2016). 자기인용 비율이 논문의 영향력을 얼마나 좌우하는지에 대한 연구는 아직도 활발히 이루어지고 있다. 예를 들어, 자기인용 비율이 본인 연구에 대한 광고도 될 수 있기 때문에 표본논문의 인용수에 긍정적인 효과가 있다는 주장이 있다(Waltman 2016). h-index나 g-index 같은 논문의 영향력 지수가 저자의 자기인용 비율에 따라 얼마나 좌우되며, 이를 통해 저자의 영향력을 예측하는 방법도 제시된다(Schreiber 2007). 본 연구에서는 자기인용 비율이 학술대회 논문의 인용수와 어떤 관련성이 있는지 분석한다.

## 4. 표본논문 수집의 절차 및 출처

학술대회 논문의 인용패턴과 그 패턴이 인용수에 끼치는 영향을 알아보는 본 연구의 목적에 맞춰 본 장에서는 이를 정보를 수집한 방법에 대해 설명한다.

### 4.1 조사 대상 표본논문의 선택 절차

본 연구는 정보과학 연구를 선도하는 국제학술대회에서 발표된 논문을 표본으로 삼았다. 즉, ACM이나 IEEE 같은 국제학술조직에서 후원하거나 주최한 국제학술대회이며 발표된 논문의 서술 언어나 구술 발표 언어가 영어인 논문을 표본논문으로 삼았다. 학술지 논문의 보편적 평가 기준인 Impact Factor에 비견될 만한 국제학술대회 평가자료가 아직 없기 때문에 우선 두 곳의 미국대학교 문헌정보대학원<sup>3)</sup>의 교

3) 워싱턴대학교 문헌정보대학원과 피츠버그대학교 문헌정보대학원

수들이 학술대회 조직위원으로 참여한 적이 있는 학술대회를 선택했다. 그리고 충분한 인용수를 확보하기 위해 2009년부터 2012년까지 4년간 발표된 논문들을 표본논문으로 선택하였다. Lisée(Lisée, Larivière and Archambault 2008)는 컴퓨터 과학 및 정보과학 분야에서 학술대회 논문의 수명이 약 8.2년이라고 주장하였다. 정보를 수집한 시기인 2017년 6월을 기준으로 논문이 발표된 지 최장 10년(2009년에 발표된 논문들)에서 최단 5년(2012년에 발표된 논문들)으로 표본논문의 인용수가 충분히 모였다고 결론짓고 그 기간을 결정하였다. 그 결과 본 연구의 표본은 총 1,904편의 학술대회 논문이고 그 표본논문은 총 31,937개의 참고자료를 인용하였다. 표본논문이 발표된 여섯 개 학술대회명과 학술대회 및 연도별 표본논문 수는 <표 2>와 같다. 참고로 해당 학술대회들은 모두 정보과학 분야의 저명 학술대회들이다. 비록 Impact Factor와 같은 명성과 객관성은 부족하나 학술대회 평가자료들 중에서 자주 언급되는 CORE ERA순위<sup>4)</sup>에서 A\* 등급(JCDL), A 등급(CSCW, CKIM, eScience), B 등급(ECIR, EDM)을 받은 학술대회들이다.

#### 4.2 참고자료 목록과 인용지수의 출처

본 연구에서 분석한 표본논문 정보, 참고자료 목록과 관련 상세 정보, 그리고 인용지수는 모두 Scopus 데이터베이스에서 수집하였다. Harzing과 Alakangas(Harzing and Alakangas 2016)는 Scopus 인용지수를 Web of Science (WoS) 인용지수와 종적으로(longitudinal) 그리고 교차학제 간으로(cross-disciplinary) 비교하였다. 그 결과 Scopus데이터베이스에 색인된 학술논문이 WoS에 비해 월등히 많아서 Scopus 인용 계수에 포함된 대상도 WoS보다 두 배 이상 많았다. 또한 2013년의 Scopus통계에 의하면 Scopus 데이터베이스에 색인된 컴퓨터 및 정보과학 분야 논문의 62.3%가 학술대회 논문이고 32.8%가 학술지 논문이었다(Scopus 2013). 그리고 같은 표본논문에 대하여 Scopus의 인용 수가 WoS보다 30% 많았다(Meho and Yang 2007). 즉 Scopus가 학술지 정보뿐만 아니라 학술대회 논문집(Conference Proceedings) 정보도 광범위하게 보유하고 있기 때문에 표본논문뿐만 아니라 참고자료의 상세정보도 안정적으로 수집할 수 있었다. 표본논문의 정보는 제목,

<표 2> 표본논문이 발표된 학술대회 목록

| 학술대회 이름  | 발표된 논문 수(학술대회 년도)                          |
|--|--|
| ACM Conference on Information and Knowledge Management, CIKM | 332 (2009), 422 (2011)                     |
| ACM Conference on Computer Supported Cooperative Work, CSCW  | 58 (2010), 112 (2011), 266 (2012)          |
| European Conference on Information Retrieval, ECIR           | 91 (2009), 72 (2010), 72 (2012)            |
| International Conference on Educational Data Mining, EDM     | 33 (2009), 52 (2010), 67 (2011)            |
| IEEE International Conference on e-Science, e-Science        | 51 (2009), 44 (2010), 54 (2011), 79 (2012) |
| ACM/IEEE-CS Joint Conference on Digital Libraries, JCDL      | 99 (2012)                                  |

(오른쪽 칸에서 팔호안의 숫자는 해당 학술대회의 년도임)

4) Core Conference Ranking Portal. <<http://www.core.edu.au/documents/>>

저자명, 참고자료 목록을 수집하였다. 참고자료의 정보는 제목, 자료의 종류, 저자명, DOI, 출간년도, 출간지 명 및 권호수, 페이지 수 그리고 URL을 수집하였다. 해당 각 자료가 참고자료로 피인용될 당시의 인용수를 파악하기 위해 각 참고자료를 인용한 논문들과 출간년도 또한 수집하였다. 그리고 저자의 자기인용 패턴을 파악하고자 표본논문 저자들이 쓴 논문들도 추가 수집하였다.

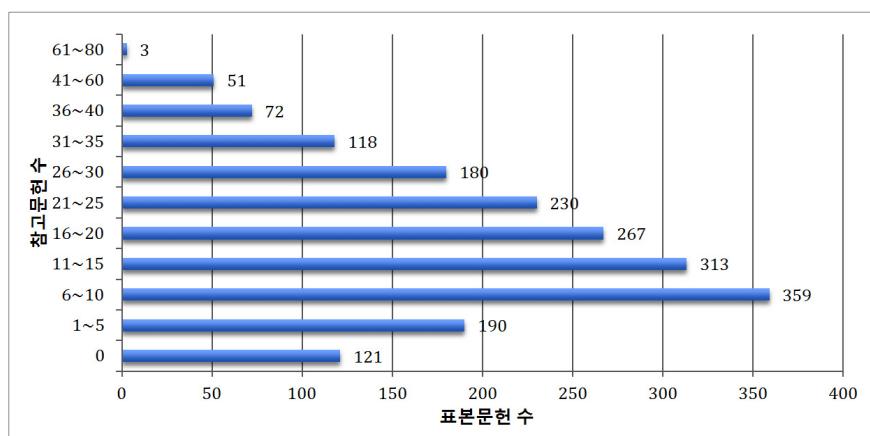
## 5. 분석 결과

본 연구의 분석은 두 단계로 진행하였다. 첫 번째 단계에서는 참고자료의 일반적인 인용패턴을 살펴보았다. 다시 말해, 논문별 참고자료의 수, 참고자료의 유형 및 나이, 참고자료가 표본논문에 피인용될 당시까지의 인용수, 자기인용 비율을 분석하였다. 두 번째 단계에서는 여러 참고자료의 특성이 표본논문의 인용수에 끼치는 영향을 회귀분석을 통해 살펴보았다.

### 5.1 참고자료 인용패턴 분석

수집된 1,904편의 표본논문은 총 31,937개의 참고자료를 인용하였다. 즉, 각 표본논문에는 평균 17.91편의 참고자료가 실렸다( $\sigma = 11.0$ ). 표본논문 중 참고자료가 전혀 없는 논문이 121 편이었고 최고 80개의 참고자료를 인용한 논문이 한 편 있었다. <그림 1>은 참고자료수에 따른 표본논문 수의 추이를 나타낸다. 참고자료 중 6~10개의 참고자료를 인용한 표본논문이 359편(전체 표본의 20.1%)으로 가장 많았고, 11~15개의 참고자료를 인용한 논문이 313편(17.6%)으로 두번째로 많았다. 전체 표본논문인 1,900여 편 중에서 70.0%( $n = 1,129$ )가 20개 이하의 참고자료를 인용한 반면, 35개 이상의 참고자료를 가진 논문은 8.5%( $n = 151$ )이었다.

다음으로는 학술대회 논문에서 인용된 참고자료의 유형을 알아보았다. 참고자료의 유형은 데이터를 수집한 Scopus의 분류를 따랐고 Scopus에서 미분류된 참고자료는 디지털 문서 식별자



<그림 1> 참고자료 개수별 논문 수

(DOI: Document Object Identifier)가 있으면 그 식별자를 통해 수동으로 유형을 수집하였다. 식별자도 없는 자료는 미분류로 남겨놓았다. 31,937개의 참고자료 중 1,627편이 미분류였고 나머지 참고자료의 유형은 <표 3>과 같다. 참고자료의 반 이상이 학술대회 논문이었으며 학술지 논문은 전체 참고자료의 29%이었다. 전체 참고자료의 3.7%, 2.7% 그리고 5.0%가 각각 웹 페이지, 단행본, 기타자료였다. 기타자료로는 뉴스레터, 기술보고서, 신문 기사, 학위논문, 설문조사 등이 있었고, 특히와 기술규격/표준도 참고자료로 쓰였지만 그 개수는 미비했다 ( $n = 6$ ). 다시 말해, 전체 참고자료의 13.8%가 학술대회 논문, 학술지, 단행본, 리뷰 등의 전통적인 참고자료가 아닌 비전통적인 자료유형이었다. 이는 문헌정보 분야에서 약 60%의 참고자료가 학술지 논문과 단행본인 한국 학술지 논문의 참고자료 인용형태와 비교하여(조인숙 2007) 큰 차이였으며 특히 단행본의 피인용비율이 국내 학술지 논문의 인용패턴과 비교하여 현저히 낮았다. 다음 장에서는 비전통적인 참고자료를 많이 인용한 표본논문이 전통적인 참고자료를 많이 인용한 논문과 비교해서 인용수에 차이가 있는지 살펴본다.

본 연구의 표본은 2009년부터 2012년까지 4년간 발표된 학술대회 논문이다. 논문 심사부터 발표까지의 기간이 짧고 다양한 발표형태가 있기 때문에 최신의 연구 아이디어를 빠르게

발표할 수 있다는 점이 학술대회 논문의 장점 중 하나이다. 따라서 학술대회 논문에서 참고자료 또한 얼마나 최신의 자료를 인용하고 있는지 살펴보았다. <그림 2>는 인용될 당시 참고자료 연력을 자료 유형별로 나타낸다. 약 31,000여편의 참고자료 중 웹 페이지와 같이 연력을 알 수 없는 자료가 1,728개였다. 연력 정보를 가진 참고자료 중 인용당시 5년 이내로 출간된 자료가 전체 참고자료의 61.2%( $n = 18,476$ )이고, 특히 인용당시 3년 이내로 출간된 참고자료가 45.1%( $n = 13,611$ )였다. 7년 이내 혹은 10년 이내로 출간된 참고자료는 각각 72.2%( $n = 21,806$ )와 83.1%( $n = 25,115$ )이었다. <그림 2>에서 볼 수 있듯이 최신 참고자료일수록 학술대회 논문의 비중이 커지고 참고자료가 오래전에 발표된 것일수록 학술지 논문의 비중이 커졌다. 예를 들어 발표된 지 2년된 참고자료의 72%가 학술대회 논문이고 22%가 학술지 논문인 반면, 발표된 지 10년 이상 지난 참고자료의 39.3%가 학술대회 논문이고 48.5%가 학술지 논문이었다. 이는 Lisée의 연구결과(Lisée, Larivière and Archambault 2008)와 마찬가지로 학술지 논문에 비해 학술대회 논문의 사장(死藏) 속도가 빠르다는 사실을 보여준다. 그리고 최근 1년 이내의 최신자료에는 웹페이지와 비전통적인 형태인 기타자료의 비중이 높았다. 발표된 지 1년 이내인 참고자료 중 웹페이지의 비중이 12.5%인 반면 발표된 지 5~10년된 참고

&lt;표 3&gt; 참고자료의 자료 유형

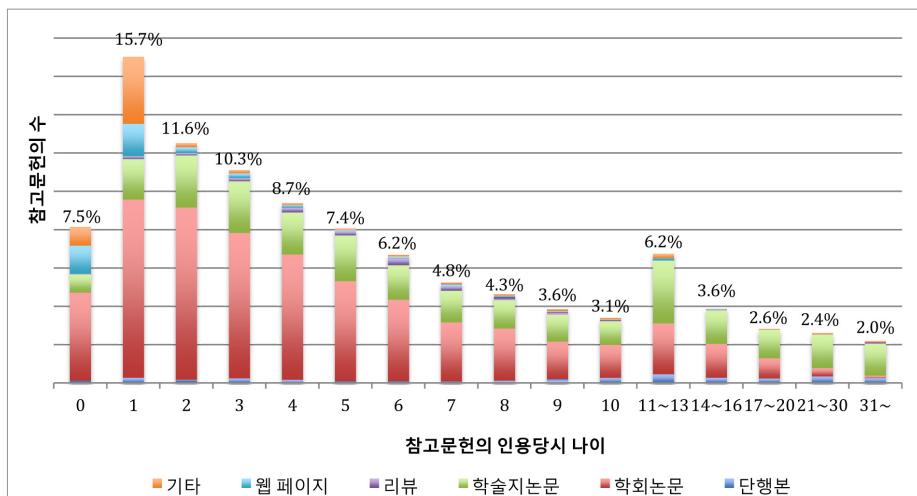
| 유형             | 학술대회 논문           | 학술지 논문           | 웹 페이지           | 단행본           | 리뷰            | 기타              | 미분류             |
|----------------|-------------------|------------------|-----------------|---------------|---------------|-----------------|-----------------|
| 참고자료 수<br>(비율) | 16,995<br>(53.2%) | 9,142<br>(28.6%) | 1,185<br>(3.7%) | 865<br>(2.7%) | 542<br>(1.7%) | 1,581<br>(5.0%) | 1,627<br>(5.1%) |

자료 중 웹페이지는 전체 참고자료의 1.1%였다. 발표된 지 1년 이내의 자료 중 기타 자료의 비중은 17.5%이고 5~10년된 참고자료 중 기타자료는 1.7%였다. 이로써 학술대회 논문에서는 최신의 정보를 얻을 때 학술적 자료(예를 들어, 학술지 논문과 학술대회 논문)뿐만 아니라 인터넷이나 기술보고서, 신문기사 등도 많이 참조하는 경향이 있었다.

다음으로 발표된 이후부터 폐인용될 때까지 해당 자료가 받은 인용수를 분석하였다. 우선 참고자료의 나이와 인용수간에는 유의미한 상관관계가 있었다( $p < .001$ ). 하지만 그 상관정도가 크지 않았다( $r = 0.28$ ). 즉, 참고자료가 발표된 지 오래되었다고 반드시 많이 이용되는 것은 아니었다. <표 4>는 참고자료 유형별 평균 인

용수를 나타낸다. 여섯가지의 유형 중 가장 인용수가 큰 참고자료는 단행본이었고 다음은 학술지 논문과 리뷰였다. ANOVA 결과에 따르면 단행본의 인용수는 다른 다섯 가지 유형에 비해 유의미하게 가장 커고 학술지 논문과 리뷰는 두 번째로 많이 인용되었고 이 인용수는 나머지 세 유형보다 유의미하게 커다. 학술지 논문과 리뷰의 인용수 간에는 통계적으로 차이가 없었다. 비록 학술대회 논문의 평균 인용수가 가장 적었지만 ANOVA 결과에 따르면 학술대회 논문, 웹페이지, 기타 자료간에는 평균 인용수에 통계적인 차이가 없었다.

마지막으로 표본논문의 저자 중에서 본인의 과거 자료를 참고자료로 인용한 자기 인용 비율을 알아봤다. 주저자이든 공저자이든 상관없이



<그림 2> 참고자료의 나이별 표본논문의 수

<표 4> 참고자료의 유형별 평균인용수

| 유형     | 학술대회 논문                       | 학술지 논문                         | 웹 페이지                          | 단행본                              | 리뷰                             | 기타                             |
|--------|-------------------------------|--------------------------------|--------------------------------|----------------------------------|--------------------------------|--------------------------------|
| 평균 인용수 | 211.1<br>( $\sigma = 725.8$ ) | 973.8<br>( $\sigma = 3138.1$ ) | 353.1<br>( $\sigma = 2385.2$ ) | 6150.5<br>( $\sigma = 10050.9$ ) | 752.5<br>( $\sigma = 1630.3$ ) | 432.9<br>( $\sigma = 2054.4$ ) |

저자 중 한 명의 자료가 참고자료에 포함된 경우 자기 인용으로 계산하였다. 전체 참고자료 중 1.65%, 즉 526개의 참고자료가 표본논문의 저자가 작성한 자료이었다. 그리고 자기 인용된 자료는 출간된 지 평균 4.53년( $\sigma = 3.83$ )이 지난 뒤 453편이 학술대회 논문이고 70편이 학술지 논문, 3편이 리뷰였다. 다음 장에서는 자기 인용된 참고자료를 가진 표본이 그렇지 않은 표본에 비해 인용수에 차이가 있는지 살펴본다.

### 5.2 인용 참고자료의 특성이 표본논문의 인용 수에 끼치는 영향력

본 장에서는 앞 장에서 분석한 다양한 특성이 그 자료를 참조한 표본논문이 앞으로 얻게 될 인용수에 어떤 영향을 끼치는지 살펴본다. 따라서 표본논문의 인용수를 종속변수로, 참고자료의 특성을 독립변수로 정의한다. 1,904편의 표본논문이 발표된 연도가 2009년부터 2012년 까지 다양하므로 발표 후 5년째 인용수를 종속 변수로 정한다. 예를 들어, 2009년에 발표된 논문은 5년간인 2009년부터 2014년까지의 인용수, 2010년에 발표된 논문은 5년간인 2010년부터 2015년까지의 인용수를 계산하여 종속변수로 사용하였다. 본 분석에서는 다음의 열 가지 특성을 독립변수로 정의했다. 아래의 목록에서 팔호안의 코드는 각 특성의 요약코드로 분석결과에서 사용된다.

- 각 표본논문의 참고자료 수(noRef)
- 각 표본논문의 참고자료 유형별 개수:
  - 학술대회 논문의 수(noConf), 학술 논문의 수(noJournal), 웹페이지의 수(noUrl), 단

행본의 수(noBook), 리뷰의 수(noReview), 기타자료의 수(noEtc)

- 각 표본논문에서 피인용될 당시 참고자료의 평균 나이(avgAge)
- 각 표본논문에서 피인용될 당시 참고자료의 평균 인용수(avgRefCite)
- 각 표본논문의 참고자료 중 자기 인용된 자료의 비율(selfCite)

본 연구에서는 표본논문의 인용수에 영향을 끼치는 피인용 참고자료의 특성을 알아보기 위해 다중회귀분석을 실시하였다. 회귀분석 실행 전에 독립 변수들은 그 레인지(Range)와 분산 값이 서로 다르므로 이를 표준화하기 위해 Z-값으로 변환하였다(Keith 2014, 541). 그리고 다중회귀분석 실행에 필요한 가설들을 검증하였다(Keith 2014, 182-212). 우선, 눈에 띄는 극단치(outlier)는 없었고 Scopus 인용지수의 표준예측치(standardized predicted values) 대비 표준잔차의 산점도(scatter plot)에서 모두 점들이 수평으로 균일하게 분포되어 있었다. 곡선을 이루는 패턴을 보이지 않았으므로 선형성 가설(assumption of linearity)도 충족되었다. 같은 산점도에서 점들의 분포가 균일하였으므로 등분산성(homoscedasticity assumption)도 충족되었다. 마지막으로 Coefficient 표에 다르면 분산팽창요인(VIF, variance inflation factor)이 참고자료의 수(noRef)만 빼고 모두 1.6 미만이었다. 참고자료의 수(noRef)는 학술대회 논문(noConf), 학술지 논문(noJournal), 기타 논문의 수(noEtc)와 최소 0.68에서 최대 0.81의 높은 상관계수를 보였으며 VIF 값도 통상적으로 허용하는 한계값인 10을 넘었으므로(Cohen et al.

2013, 423-424) 회귀분석에서 제외하였다. 그 외 아홉 가지 독립변수 중 최소 VIF값은 1.03이었고 최고 VIF값은 1.22였다. 이로써 다중공선성의 가설(assumption of the multicollinearity)은 더 이상 문제가 없었다(Keith 2014, 182-212).

모든 관련 가설이 충족되므로 다중회귀분석을 실시하였다. 그 결과는 〈표 5〉와 같다. 아홉 개의 참고자료 관련 특성으로 구성된 회귀식은 표본논문의 인용수 변화 중 8.1%를 설명한다 ( $F = 18.59, p < .001; R^2 = .081$ , 수정  $R^2 = .077$ ). 0.05 유의수준에서 참고자료로 쓰인 학술대회 논문의 수(*noConf*:  $\beta = 4.95; sr^2 = .05$ ), 기타자료의 수(*noEtc*:  $\beta = 1.95; sr^2 = .01$ ) 그리고 참고자료의 평균나이(*avgAge*:  $\beta = -1.06; sr^2 = .002$ )가 인용 수에 통계적으로 유의미한 영향력을 가졌다. 이 중 학술대회 논문의 수가 해당 표본논문이 받게 될 인용수를 가장 크게 좌우했고 기타자료의 수는 그 다음 수준으로 영향이 있었다. 두 영향력은 모두 양의 값이다. 다시 말해, 학술대회 논문이나 기타자료가 참고자료로 많이 쓰일수록 해당 표본논문이 많이 인용되었다. 하지만 전통적으로 참고자료로서

많이 인용되는 학술지논문이나 리뷰를 많이 참조하더라도 표본논문의 인용수에는 영향이 없었다. 그리고 참고자료의 나이도 유의미한 영향력이 있었는데 음의 영향력이었다. 즉, 최신의 자료를 참고자료로 많이 쓸수록 해당 표본논문도 많이 인용되었다. 이는 참고자료의 최신성이 표본논문의 인용수를 좌우하는 중요한 요소라는 기존 연구(Onodera and Yoshikane 2015)와 같은 결과이다. 하지만 피인용될 당시 자료의 명성을 나타내는 참고자료의 인용수는 표본논문의 인용수에 전혀 영향을 끼치지 않았다. 즉, 저명한 논문을 참고자료로 쓰더라도 표본논문의 인용수에는 변화가 없었다. 이는 나노과학분야로 해당분야는 다르지만 기존 연구(Didegah and Thelwall 2013)의 결과와 상충되는 것으로 추후에 보다 심도 있는 연구가 필요하다. 그리고 기존 논문들(Schreiber 2007)은 참고자료의 자기 인용비율이 높을수록 본인 연구의 광고효과로 표본논문의 인용수도 같이 높아진다고 주장하였으나 본 연구에서는 자기 인용비율이 표본논문의 인용수에 끼치는 영향이 전혀 없었다.

〈표 5〉 회귀분석 결과

| 변수                | B       | SEB | $\beta$ | $sr^2$ |
|-------------------|---------|-----|---------|--------|
| <i>noConf</i>     | 4.95 *  | .50 | .23     | .05    |
| <i>noJournal</i>  | .47     | .58 | .02     | .00    |
| <i>noURL</i>      | -.74    | .47 | -.04    | .00    |
| <i>noBook</i>     | .08     | .52 | .00     | .00    |
| <i>noReview</i>   | -.11    | .48 | -.01    | .00    |
| <i>noEtc</i>      | 1.95 *  | .53 | .09     | .01    |
| <i>avgAge</i>     | -1.06 * | .52 | -.05    | .00    |
| <i>avgRefCite</i> | .90     | .51 | .04     | .00    |
| <i>selfCite</i>   | .63     | .48 | .03     | .00    |

(별표\*는 통계적으로 유의미한 영향력을 나타냄.  $p < .05$ )

## 6. 결 론

본 연구는 정보과학 분야를 중심으로 여러 참고자료가 어떻게 인용되는지 패턴을 살펴보고 학술대회 논문의 인용 수에 끼치는 영향이 무엇인지 알아보았다. 참고자료의 피인용 패턴이나 표본논문의 인용수에 영향을 끼치는 여러 특성에 대한 기존 연구는 학술지 논문을 중심으로 활발히 진행되고 있다. 하지만 정보과학 연구에서 학술대회 논문의 비중이 점점 커지고 있음에도 학술대회 논문을 대상으로 하는 연구는 아직 초기 단계에 머물러 있다는 점에서도 학술대회 논문을 대상으로 한 본 연구의 의미가 있다. 본 논문은 연구의 중요 지식 및 배경 자료로써 쓰이는 참고자료가 어떻게 인용되는지에 초점을 맞추었다. 관련 참고자료의 특성으로, 참고자료의 수, 참고자료의 유형, 참고자료로 쓰일 당시 참고자료의 인용수 및 나이, 저자의 자기인용 비율을 조사하였다.

그 결과, 표본논문은 평균 18편의 참고자료를 인용하였고 표본의 70%가 20개 이하의 참고자료를 인용하였다. 하지만 다중회귀분석에서는 높은 다중공선성으로 참고자료의 수가 표본의 인용수에 끼치는 영향을 분석할 수 없었다. 참고자료의 유형에 대한 분석으로는 피인용 참고자료의 53%가 학술대회 논문이었고, 29%가 학술지 논문이었다. 특히, 참고자료의 14%가 논문이나 단행본이 아닌 비전형적인 참고자료의

유형이었다. 이후 실행한 다중 회귀분석 결과에 따르면 학술대회 논문을 참고자료로 많이 쓴 표본논문일수록 많이 인용되었고 비전형적인 참고자료 중 기타자료를 참고자료로 많이 쓴 표본논문도 인용수가 높았다. 전체 참고자료의 60% 이상이 5년 이내에 발표된 자료이고, 45%는 3년 이내에 출간된 자료이었다. 그리고 최신 참고자료일수록 학술대회 논문, 웹 페이지와 기타 자료의 비중이 높았다. 다중회귀분석 결과에 따르면 최신의 자료를 많이 참조한 표본논문일수록 추후 인용수를 많이 얻었다. 하지만 참고자료의 명성을 나타내는 참고자료의 인용수는 표본논문이 얻게 될 인용수에 크게 영향이 없었다. 즉, 많이 인용되어 이미 명성을 얻은 자료를 참고자료로 쓰더라도 표본논문의 인용수에는 변화가 없었다. 자기 인용 비율도 1.7%로 미비했고 표본논문의 인용수에 유의미한 영향을 끼치지 않았다.

추후 연구 과제로는 보다 많은 표본논문을 수집하고 학술대회에 발표된 논문의 저자, 참고자료, 논문 자체의 내용적 특징 등 논문 자체의 세부 특징뿐만 아니라 발표된 학술대회의 특성이 해당 논문에 어떤 영향을 끼치는지 살펴볼 것이다. 또한 Google Scholar를 포함해 다른 색인 데이터베이스에서도 인용수를 수집하여 학술대회 논문의 인용수 자체에 대한 연구도 할 것이다.

## 참 고 문 헌

- [1] 김갑선. 2002. 참고문헌의 동시공존현상-한국 사회과학자들의 인용동기와 참고자료의 분석. 『한국 문헌정보학회지』, 36(4): 21-47.
- [2] 김미라, 이응봉. 2013. 참고문헌 분석을 통한 국방과학 연구자들의 인용행태 연구. 『한국문헌정보학 회지』, 47(2): 143-159.
- [3] 오세훈. 2005. 우리나라 문헌정보학 학술지 논문 및 인용문헌 분석을 통한 연구동향 연구. 『정보관 리학회지』, 22(3): 379-408.
- [4] 조인숙, 한미숙. 2007. 문헌정보학분야 연구자들의 학술지 게재논문과 인용학술지 분석 연구. 『한국 비블리아학회지』, 18(2): 89-105.
- [5] Antoniou, G. A. et al. 2015. "Bibliometric Analysis of Factors Predicting Increased Citations in the Vascular and Endovascular Literature." *Annals of Vascular Surgery*, 29(2): 286-292.
- [6] Barbosa, S. D. J., Silveira, M. S. and Gasparini, I. 2017. "What Publications Metadata Tell Us about the Evolution of a Scientific Community: The Case of the Brazilian Human-Computer Interaction Conference Series." *Scientometrics*, 110(1): 275-300.
- [7] Bartneck, C. and Hu, J. 2009. "Scientometric Analysis of the CHI Proceedings." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, April 04-09, 2009, Boston, MA, USA: 699-708.
- [8] Biscaro, C. and Giupponi, C. 2014. "Co-Authorship and Bibliographic Coupling Network Effects on Citations." *PLOS ONE*, 9(6): e99502.
- [9] Bornmann, L. et al. 2012. "Diversity, Value and Limitations of the Journal Impact Factor and Alternative Metrics." *Rheumatology International*, 32(7): 1861-1867.
- [10] Chakraborty, T. et al. 2014. "Towards a Stratified Learning Approach to Predict Future Citation Counts." In *Proceedings of the 14th ACM/IEEE Joint Conference on Digital Libraries*, September 08-12, 2014: 351-360.
- [11] Cohen, J. et al. 2013. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences*. New York: Routledge.
- [12] Didegah, F. and Thelwall, M. 2013. "Determinants of Research Citation Impact in Nanoscience and Nanotechnology." *Journal of the American Society for Information Science and Technology*, 64(5): 1055-1064.
- [13] Donner, P. 2016. "Enhanced Self-citation Detection by Fuzzy Author Name Matching and Complementary Error Estimates." *Journal of the Association for Information Science and*

- Technology*, 67(3): 662-670.
- [14] Eckmann, M., Rocha, A. and Wainer, J. 2012. "Relationship Between High-Quality Journals and Conferences in Computer Vision." *Scientometrics*, 90(2): 617-630.
  - [15] Falagas, M. E. et al. 2013. "The Impact of Article Length on the Number of Future Citations: A Bibliometric Analysis of General Medicine Journals." *PLoS One*, 8(2): e49476.
  - [16] Franceschet, M. 2010. "A Comparison of Bibliometric Indicators for Computer Science Scholars and Journals on Web of Science and Google Scholar." *Scientometrics*, 83(1): 243-258.
  - [17] Freyne, J. et al. 2010. "Relative Status of Journal and Conference Publications in Computer Science." *Communications of the ACM*, 53(11): 124-132.
  - [18] Harzing, A.-W. and Alakangas, S. 2016. "Google Scholar, Scopus and the Web of Science: A Longitudinal and Cross-Disciplinary Comparison." *Scientometrics*, 106(2): 787-804.
  - [19] Keith, T. Z. 2014. *Multiple Regression and Beyond: An Introduction to Multiple Regression and Structural Equation Modeling*. New York: Routledge.
  - [20] Larivière, V., Sugimoto, C. R. and Cronin, B. 2012. "A Bibliometric Chronicling of Library and Information Science's First Hundred Years." *Journal of the American Society for Information Science and Technology*, 63(5): 997-1016.
  - [21] Lisée, C., Larivière, V. and Archambault, É. 2008. "Conference Proceedings as A Source of Scientific Information: A Bibliometric Analysis." *Journal of the American Society for Information Science and Technology*, 59(11): 1776-1784.
  - [22] Meho, L. I. and Yang, K. 2007. "Impact of Data Sources on Citation Counts and Rankings of LIS Faculty: Web of Science Versus Scopus and Google Scholar." *Journal of the American Society for Information Science and Technology*, 58(13): 2105-2125.
  - [23] Montesi, M. and Owen, J. M. 2008. "From Conference to Journal Publication: How Conference Papers in Software Engineering are Extended for Publication in Journals." *Journal of the American Society for Information Science and Technology*, 59(5): 816-829.
  - [24] Nederhof, M. J. 2006. "Bibliometric Monitoring of Research Performance in the Social Sciences and the Humanities: A Review." *Scientometrics*, 66(1): 81-100.
  - [25] Onodera, N. and Yoshikane, F. 2015. "Factors Affecting Citation Rates of Research Articles." *Journal of the Association for Information Science and Technology*, 66(4): 739-764.
  - [26] Sakr, S. and Alomari, M. 2012. "A Decade of Database Conferences: A Look Inside the Program Committees." *Scientometrics*, 91(1): 173-184.
  - [27] Schreiber, M. 2007. "Self-Citation Corrections for the Hirsch Index." *EPL (Europhysics Letters)*, 78(3): 30002.

- [28] Scopus. 2013. Scopus Content Coverage Guide. [online] [cited 2017. 9. 20.]  
⟨<https://files.sciverse.com/documents/pdf/ContentCoverageGuide-jan-2013.pdf>⟩
- [29] Shirakawa, N. et al. 2012. "Global Competition and Technological Transition in Electrical, Electronic, Information and Communication Engineering: Quantitative Analysis of Periodicals and Conference Proceedings of the IEEE." *Scientometrics*, 91(3): 895-910.
- [30] Tahamtan, I., Safipour Afshar, A. and Ahamdzadeh, K. 2016. "Factors Affecting Number of Citations: A Comprehensive Review of the Literature." *Scientometrics*, 107(3): 1195-1225.
- [31] van Wesel, M., Wyatt, S. and ten Haaf, J. 2014. "What a Difference a Colon Makes: How Superficial Factors Influence Subsequent Citation." *Scientometrics*, 98(3): 1601-1615.
- [32] Vardi, M. Y. 2010. "Revisiting the Publication Culture in Computing Research." *Communications of the ACM*, 53(3): 5.
- [33] Vasilescu, B. et al. 2014. "How Healthy are Software Engineering Conferences?" *Science of Computer Programming*, 89(C): 251-272.
- [34] Vrettas, G. and Sanderson, M. 2015. "Conferences Versus Journals in Computer Science." *Journal of the Association for Information Science and Technology*, 66(12): 2674-2684.
- [35] Wainer, J. and Valle, E. 2013. "What Happens to Computer Science Research after It Is Published? Tracking CS Research Lines." *Journal of the American Society for Information Science and Technology*, 64(6): 1104-1111.
- [36] Wainer, J., Przibiszczki de Oliveira, H. and Anido, R. 2011. "Patterns of Bibliographic References in the ACM Published Papers." *Information Processing & Management*, 47(1): 135-142.
- [37] Waltman, L. 2016. "A Review of the Literature on Citation Impact Indicators." *Journal of Informetrics*, 10(2): 365-391.
- [38] Wuehrer, G. and Smejkal, A. 2013. "The Knowledge Domain of the Academy of International Business Studies (AIB) Conferences: A Longitudinal Scientometric Perspective for the Years 2006-2011." *Scientometrics*, 95(2): 541-561.
- [39] Zhang, Y. and Jia, X. 2013. "Republication of Conference Papers in Journals?" *Learned Publishing*, 26(3): 189-196.

• 국문 참고자료의 영어 표기

(English translation / romanization of references originally written in Korean)

- [1] Kim, Kapseon. 2002. "An Analysis on the Citation Motivations and References by Social Scientists: The Co-occurrence Phenomenon of Both Korean and Non-Korean Literatures

- Within the Korean Reference." *Journal of the Korean Society for Library and Information Science*, 36(4): 21-47.
- [2] Kim, Mi-Ra and Lee, Eung-Bong. 2013. "A Study on the Citing Behavior of Military Scientists by Reference Analysis." *Journal of the Korean Society for Library and Information Science*, 47(2): 143-159.
- [3] Oh, Se-Hoon. 2005. "A Study on the Research Trends of Library & Information Science in Korea by analyzing Journal articles and the Cited Literatures." *Journal of the Korean Society for Information Management*, 22(3): 379-408.
- [4] Cho, In-Sook and Han, Mi-Sook. 2007. "An Analysis on the Articles and Cited Journals Suggested by LIS Researchers of Korea." *Journal of the Korean Biblia Society for Library and Information Science*, 18(2): 89-105.