

Cited-by Linking 데이터 기반 자연과학 및 공학 분야 학술논문 인용 패턴 분석

Citing Pattern Analysis based on Cited-by Linking Data of DOI Journals in the Field of Natural Sciences & Engineering

서선경 (Sun Kyung Seo)*, 최호남 (Ho Nam Choi)**
김병규 (Byung-Kyu Kim)***, 최선희 (Seon-Heui Choi)****
김정환 (Jeong Hwan Kim)*****

초 록

Cited-by Linking 서비스는 CrossRef에서 제공하는 주요 서비스 중 하나로 해당 논문이 DOI를 통하여 얼마나 인용되었는지 추적하여 집계된 데이터이다. 이에 본 연구에서는 KISTI의 학술정보통합관리시스템에서 월단위로 구축·관리하는 Cited-by Linking 데이터를 분석하여 자연과학과 공학 분야의 인용 패턴을 규명하고자 하였다. 이를 위해서 전체 기탁된 21만 건 중 자연과학과 공학 분야 총 170,999건(315종)을 분석 대상으로 하고, 2016년 3월까지 누적된 Cited-by Linking 데이터를 수집하였다. 연구의 분석 결과, 첫째, 글로벌하게 인용될 가능성에서 사용 언어는 영향을 미치고 있으며, 둘째, SCIE 및 SCOPUS 등재 여부 역시 인용가능성에 상당 부분 기여하고 있음을 확인하였다. 셋째, 자연과학 분야는 거의 동일한 주제 분야에서 인용을 받고 있으며, 상대적으로 공학 분야는 다주제 분야에서 더 인용 받고 있음을 파악하였다. 이러한 연구의 결과는 자연과학과 공학 분야의 세부주제 분야별 인용 행태를 규명하고, 향후, 인용 행태에 관한 연구에서 유용하게 활용될 것으로 기대한다.

ABSTRACT

Cited-by Linking Service is one of the CrossRef's information services that allows you to discover how your publications are being cited and to incorporate that information into your online publication platform. This study tries to investigate citation patterns in the field of both Natural Science and Engineering using all of DOI assigned articles and Cited-by Linking data which are accumulated and managed by KISTI. The investigating approach is designed to verify the theory of 1) cognitive accessibility, 2) 'perceived quality and significance' and 3) 'subject relativity'. For cognitive accessibility verification the fulltext language portion of Korean and English between "Cited DOI Source Data" and "NOT Cited DOI Source Data" was compared. For perceived quality and significance verification the availability of the "Cited DOI Source Data" and "NOT Cited DOI Source Data" from SCIE and SCOPUS was employed. For subject relativity DOI data were classified and analysed on the basis of OECD subject classification scheme. Findings are that global citability is closely related to the fulltext language of the articles and their quality and significance. And in the natural science field most of citations are from the same subject categories, while relatively more citations are from other subject categories in the engineering field.

키워드: 인용패턴분석, 인용행태
citing pattern analysis, Cited-by Linking Data, DOI articles

-
- * 한국과학기술정보연구원 정보기반실(welloffssk@naver.com) (제1저자)
** 한국과학기술정보연구원 정보기반실(hnchoi@kisti.re.kr) (교신저자)
*** 한국과학기술정보연구원 정보기반실(bk.kim@kisti.re.kr) (공동저자)
**** 한국과학기술정보연구원 정보기반실(sunny.choi@kisti.re.kr) (공동저자)
***** 한국과학기술정보연구원 정보기반실(kimjh@kisti.re.kr) (공동저자)

- 논문접수일자: 2016년 5월 23일 ■ 최초심사일자: 2016년 5월 25일 ■ 게재확정일자: 2016년 6월 6일
■ 정보관리학회지, 33(2), 157-176, 2016. [http://dx.doi.org/10.3743/KOSIM.2016.33.2.157]

1. 서론

계량정보학 분야에서 인용 분석은 다각적 측면, 다양한 레벨에서 분석을 통해 연구가 진행되고 있다. 인용 분석을 하는 연구자들에게 “인용 횟수는 무엇을 측정하는가?”의 질문을 하였을 때, 명쾌한 해답을 얻기 위해서는 인용 분석에 대한 철학적 배경이 필요하다. 과학의 규범 이론(normative theory)을 지지하는 연구자들은 인용 분석을 통해 학술지, 논문, 연구자, 기관의 질적 측정이 가능하다는 주장을 한다. 이들은 계량정보학 분야에서 지속적으로 인용 분석을 연구하고, 인용 횟수를 집계하는 것을 시작으로 끊임없이 그 기술을 발전시키고 있다. 그러나 미시사회학적 관점의 연구자들은 인용 동기에 관한 다양한 요인을 제시하며, 규범 이론이 갖는 가정의 유효성에 대한 의문과 비판을 제기한다. 이와 같은 접근에서 이들은 인용을 질적인 측면에서 측정 가능한 지표로 삼아 품질을 평가하는 것에 대해 부정적인 견해를 표명했다.

약 15년간의 과학자들의 인용 행태에 관한 연구들을 검토한 Lutz와 Hans(2008)는 인용이란 일차원적 현상이 아닌 다차원적인 현상이며, 비과학적인 측면에서 다양한 요인들의 역할이 큰 것을 확인했다. Van Raan(2005)은 연구 기간 전체에 걸쳐 연구원 그룹의 모든 연구 성과물을 대상으로 인용 분석을 적용하면, 인용은 다양한 상황에서 과학적 성과에 대한 강한 지표가 될 수 있을 것이라 했다. Van Raan(2005)

은 과학의 인지적 내용에 관한 사회적 요인들의 영향력에서 Cole(1992)의 연구를 통하여 개인적인 연구자 단계와 인용을 위한 상위 집합적 단계로 구분하였다. Cole(1992)에 의하면 “정상 과학(normal science)”이 진행되는 단계에서 거시적 수준의 “공동 지식 성과(communal knowledge outcome)”는 과학의 규범 이론이 적절하다는 것이다. Liu(1997)는 인용이 복잡한 지적 관례임을 인정하고, 인용 빈도에 기여하는 요인들을 4개의 주요 범주로 구분하여 제시하였다. 첫째, 물리적 접근성(Physical accessibility)은 논문을 인식하지 못하거나 그것을 획득하지 못하는 경우에 그 논문을 읽을 수 없을 것이다. 둘째, 인지적 접근성(Cognitive accessibility)은 논문을 이해하는 것에 어려움이 있거나 언어의 장벽으로 읽을 수 없을 때의 영역이다. 셋째, 품질(Perceived quality)은 논문의 수준이 낮다고 인식되는 경우 덜 인용될 것이라는 것이다. 마지막으로 중요성(Perceived significance)은 만약 논문이 과학적으로 덜 중요한 경우, 물리적이거나 인지적으로 가능하여도 덜 인용될 것으로 접근하였다.

한편, KISTI는 2008년부터 과학기술분야의 학술논문 개별에 대하여 고유의 식별자(DOI: Digital Object Identifier)를 부여하고, 국제 DOI 재단(IDF: International DOI Foundation) 산하의 기관인 CrossRef¹⁾(DOI RA: Registration Agency)에 DOI를 기탁하는 서비스 사업을 진행하고 있다. DOI 서비스는 그 식별자를 이용하여 학술논문의 영속적인 접근을 보장한다.

1) CrossRef(<http://www.crossref.org>)는 적절한 기술 인프라를 활용한 상호협력증진을 통해 온라인 학술정보의 용이한 활용을 도모하는 목적으로 2000년 초 출범되었다. DOI를 등록할 수 있는 등록기관 중 하나로써 학술정보와 관련된 디지털객체의 DOI 등록업무와 DOI를 검색할 수 있는 검색시스템을 제공하는 서비스를 실시한다.

CrossRef는 등록된 DOI와 메타데이터를 이용하여 학술 커뮤니케이션에 필요한 응용서비스를 개발하고, 서비스하고 있다. 주요 서비스 중 하나인 Cited-by Linking 서비스는 해당 논문이 DOI를 통하여 얼마나 이용되었는지 알려주는 서비스이다(서태설, 최희운, 2011). DOI가 부여된 개별 논문이 CrossRef에 기탁되면 웹을 통한 끊임없는 서비스뿐만 아니라 DOI 참고문헌 링크를 통하여 노출이 더 많이 될 가능성을 지니고 있다.

이와 같은 Cited-by Linking 서비스의 데이터를 활용하여, 본 연구에서는 자연과학과 공학 분야의 데이터를 분석대상으로 다음과 같은 가설에 대한 분석을 실시한다.

첫째, 인지적 접근성의 측면에서 영문지가 국문지보다 더 인용이 많이 될 것이다. 이를 검증하기 위해 DOI 참여 학술지 중 1번이라도 인용 받은(Cited) 논문들의 사용 언어 비율과 DOI 참여 학술지 중 인용 받지 않은 논문들의 사용 언어 비율을 분석한다.

둘째, 품질과 중요성의 측면에서 SCIE와 SCOPUS 등재지의 경우 더 많이 인용 될 것이다. 이를 검증하기 위해 DOI 참여 학술지 중 1번이라도 인용 받은(Cited) 논문들의 SCIE와 SCOPUS 등재 여부 비율과 DOI 참여 학술지 중 인용 받지 않은 논문들의 SCIE와 SCOPUS 등재 여부 비율을 분석한다.

셋째, 다수의 논문들이 유사한 주제 분야에서 인용되었을 것이다. 이는 다양한 인용 동기 중 주제적 연관성을 살펴보기 위한 측면이다. 이를 검증하기 위해 DOI 참여 학술지에 속한 논문들을 인용한(Citing) 해외 또는 국내 논문들 중 SCIE 논문의 주제 분야를 분석한다.

따라서 본 연구에서는 Cited-by Linking 데이터를 분석하여, 위와 같은 3개의 가설에 대한 근거를 마련하고, 자연과학과 공학 분야의 세부주제 분야별 인용 패턴을 규명하는 것을 목적으로 한다. 이와 같은 인용 패턴 분석의 결과는 향후 자연과학과 공학 분야의 연구자들에게 자신들의 연구 성과를 더 널리 공유할 수 있게 하는 방향을 제시할 것이다. 더불어, 국내 학술지의 국제적 활성화를 위한 방안에 대한 객관적인 근거를 도출할 것으로 기대한다.

2. 관련 연구

관련 연구는 크게 인용 행태에 관한 국내, 국외의 선행연구와 DOI의 Cited-by Linking 데이터를 분석한 국내 연구를 중심으로 정리하였다. 국내에서는 조현양과 조현선(2005)이 공학 분야 가운데 4개 주요 분야에서 발행되는 학회지에 수록된 논문에 대한 인용 분석을 실시하여, 분야별 연구자의 정보 이용 행태를 파악하였다. 이 연구에서는 기계, 건축, 전기, 전자 분야의 연구자들의 최신 정보에 대한 요구 정도, 주요 정보원의 형태 및 정보원에 대한 수명 차이를 확인하기 위해 인용된 참고문헌의 유형, 논문 1편당 평균 참고문헌 건수, 연도별 건수, 언어별 건수를 분석하고, 인용된 정보의 반감기에 있어 분야별 차이를 규명하였다.

또한 장덕현과 장환석(2007)은 기계공학 연구자들의 인용 행태를 분석하기 위해 P대학교의 기계공학과에서 1996년과 2004년에 생산된 박사학위논문을 대상으로 인용 분석을 실시하고, 각 년도간의 인용 행태 차이를 조사하였다.

이 연구에서 인용 행태의 변화를 확인하기 위해 설정된 가설은 인용문헌의 유형, 언어, 출판 이후 경과시간, SCI 등재 학술지 인용 비율이었다. 이 연구는 학위논문을 대상으로 인용 분석을 시도하여 박사과정의 연구자들이 어떤 정보 이용행태를 따르는지를 확인할 수 있었으며, 과거와 현재의 비교를 통해 정보이용 행태가 향후 어떤 모습으로 변하게 될지 예측할 수 있는 기초를 마련하였다.

국내 과학기술 분야 연구자들의 해외 학술지 인용 행태에 대한 분석을 하기 위하여 김병규, 강무영, 최선희, 김수영, 유범중, 신재도(2011)는 KISTI의 KSCD(한국 과학기술 인용색인 데이터베이스)를 활용하여 해외 인용문헌의 형태별 분석, 주제별 피인용수 및 타분야 인용비율 분석, 해외 인용문헌의 언어, 발행국가, 발행기관 분석 등을 실시하였다. 이 연구는 해외 학술지 영향력에 대한 종합적 분석이 가능하였으며, 이를 바탕으로 해외 전자저널 구독 전략 수립에 대한 방향을 제시에 의의가 있다.

국방 과학 연구자들의 인용 행태를 연구하기 위하여 김미라, 이응봉(2013)은 국방과학 분야 국내 대표 학술지에 게재된 학술지 논문의 참고문헌을 분석하였다. 조사 기간은 1998년~2012년으로 한국군사과학기술학회지 총 537편의 논문에 인용된 참고문헌 4,563건에 관해서 논문이 인용된 평균 참고문헌 수 및 참고문헌의 증가량, 내부생산물인 보고서와 외부생산물인 학회지 논문과의 평균 참고문헌 수 및 자료유형별 선호도 차이, 주제 분야에 따른 자료 유형별 선호도 차이가 있음을 비교·분석하였다. 이는 연구도서관이 속한 모기관으로부터 예산 확보의 기초자료로 활용될 수 있다는 발판을 마련

하였다.

국외의 경우에는 국내보다 인용 행태에 관한 연구가 활발하게 시도되고 있다. Gross와 Gross(1927)에 의해 최초로 인용 분석이 시작되었으며, 이들은 화학 분야의 대표 학술지를 선정하여 참고문헌 3,633건을 대상으로 화학 분야의 학술 문헌의 중요도를 측정하였다. 이를 시작으로 각 주제 분야별로 연구자들의 정보 이용 패턴을 규명하기 위해 계량정보학자들의 인용 분석 연구가 다각적으로 이루어지고 있다.

Davis와 Cohen(2001)은 코넬대학교 학부생의 1996년과 1999년의 학기말 레포트를 대상으로 인용 분석을 실시하여, 참고문헌 유형이 단행본에서 신문과 웹자원으로 이동하고 있음을 확인하였다. 단행본 인용은 30%에서 19%로 감소하고, 신문은 7%에서 19%로, 웹자원은 9%에서 21%로 증가하고 있음을 규명함에 따라 온라인 데이터베이스의 증가로 인하여 신문과 웹자원 인용 비율이 꾸준히 증가할 것을 예상하였다.

Lutz와 Hans(2008)는 지난 15년 동안 발표된 과학자들의 인용 행태에 관한 연구들을 전체적으로 검토하였다. 이 연구의 분석들을 바탕으로 과학자들이 인용하는 동기에 관하여 답을 찾으려고 시도하였으며, 과학자들의 동료 간의 지적 인정과 인지적인 영향력뿐만 아니라 비과학적인 이유들도 포함될 가능성이 있음을 정리하였다.

Obuh와 Babatope(2011)는 문헌정보학과 물리학 분야의 학생의 인용 행태를 분석하기 위하여, 델타 주립대학 학부생들의 프로젝트 문헌 50건을 대상으로 그에 속하는 참고문헌의 유형과 포맷을 조사하였다. 유형에서는 문헌정보학부생들이 단행본을, 물리학부생들은 학술지를

더 선호하였으며, 포맷은 인쇄자료와 전자자료 중 문헌정보학부생들은 53 대 47로 전자자료 대 인쇄자료를, 물리학부생들은 77 대 23으로 전자자료 대 인쇄자료를 선호하고 있는 것으로 규명하였다.

Milojevic'(2012)은 천문학, 수학, 로봇공학, 생태학, 경제학 분야의 핵심 학술지 논문 213,756건을 조사하여 인용 행태를 분석하였다. 지난 50년 동안의 참고문헌 이용과 특징을 각 분야별로 비교하여 결과를 제시하였을 때, 5개 분야 모두 논문 당 참고문헌의 비율이 증가하고 있었으며, 1990년 말에 인터넷의 영향으로 참고문헌의 비율이 변화될 것을 예측하였지만, 이는 확인되지 않았다. 이들은 논문 당 참고문헌의 비율이 증가한 요인을 2가지로 제시하였으며, 이는 논문의 길이가 길어진 것과 참조의 비율이 높아진 것으로 정리하였다.

이상의 선행연구들에서 살펴본 것과 같이 인용 행태 분석은 주제분야별, 분석 대상별로 국내·외에서 다양하게 진행되고 있다. 이와 같이 인용 행태 분석은 그 분야의 정보 인용 패턴을 확보하고, 이를 통해 미래의 정보 채널을 예측하거나 도서관 또는 연구기관에서 예산 마련에 관한 기초 정보를 제공해 줄 수 있을 것이다.

DOI의 Cited-by Linking 데이터를 분석한 국내 연구는 정은경, 김병규, 전지영, 최선희, 강무영(2011)이 Cited-by Linking 서비스를 소개하며, 2009년 11월, 2010년 2월, 2010년 7월의 기탁된 논문을 대상으로 1회 이상 인용된 논문의 시점별 인용 추이를 분석하는 것을 시작되었다. 이 연구에서는 1회 이상 인용된 논문이 증가하는 추세이며, 1개 논문 최대 피인용수 역시 증가함을 확인하여 DOI 번호를 부여한 논문의

기탁은 국제 유통망에서 국내 학술지 논문의 노출 빈도를 높일 수 있을 것으로 기대했다.

Seo, Jung, & Kim(2013)은 Cited-by Linking 데이터와 국내 학술데이터베이스(KSCD)를 활용하여 국내 DOI 참여 학술지의 인용 패턴을 분석하였다. 분석 대상은 Cited-by Linking과 KSCD의 비교를 위해 2002년부터 2010년까지의 68종 학술지의 53,147건 논문을 수집하였다. 이 연구에서는 DOI 참여 학술지를 인용한 학술지의 국가와 언어를 분석하고, 가장 많이 인용한 해외 학술지 20위를 제시하였다. 또한 Cited-by Linking을 통해 1회 이상 인용 받은 논문 건수가 KJCR의 1회 이상 인용 받은 논문 건수보다 더 많은 것을 확인하여, 국제적으로 인용될 기회가 더 많다는 것을 언급하였다. 그리고 기탁된 DOI 논문 중 국문으로 작성된 논문이 국내에서 인용을 받은 비율이 높고, 영문을 작성된 논문이 해외에서 인용을 받은 비율이 높음을 확인하였으며, 더불어 SCI(E) 학술지 8종과 SCI(E)가 아닌 학술지 60종의 인용률을 비교하여 SCI(E)의 학술지가 인용률이 더 높음을 확인하였다.

Cited-by Linking 데이터는 CrossRef를 통해 국제적으로 수집되기 때문에 WoS, SCOPUS, KSCD보다 커버리지가 더 광범위할 것이다. 앞으로 이를 활용하여 시도되는 국내 인용 분석 연구들이 활발하게 진행될 것으로 예상된다. 본 연구에서는 Cited-by Linking 데이터를 활용하여, 국내 자연과학과 공학 분야의 DOI 논문의 1번 이상 인용된 논문과 1번도 인용되지 않은 논문의 사용 언어 구성비(국문 대 영문), SCIE 등재 여부 비, SCOPUS 등재 여부 비를 비교·분석하고자 한다. 뿐만 아니라, 국내 자연

과학과 공학 분야의 DOI 논문을 인용한 SCIE 논문의 주제 분야를 분석하여 세부주제 분야별 인용 행태를 규명한다.

3. Cited-by Linking 데이터 분석 과정

3.1 DOI 기탁 현황

인터넷과 정보기술의 발달은 학술정보 유통의 경로를 변화시켰다. 학술정보는 비전자적인 종이 인쇄물(Physical Object) 또는 디지털 객체(Digital Object) 형태의 데이터 객체(Data Object)로써 출판·유통되고 있다. 이에 학술정보를 인터넷에서 고유하게 식별하여 항구적인 접근을 보장하기 위해 DOI 개념이 제안되었다. DOI는 URL(Uniform Resource Locator)과 유사한 개념이지만 하나의 객체가 다른 서버로 옮겨지면 변하는 URL과는 다르게 한번 생성된 DOI는 시간과 장소에 제한을 받지 않는 장점이 있다.

KISTI에서는 CrossRef와의 협력을 바탕으로 국내 과학기술분야의 학술논문에 CrossRef DOI 번호를 부여하고, 논문과 그 논문의 참고

문헌 메타데이터를 기탁하는 DOI 사업을 수행해오고 있다. CrossRef의 제공 서비스는 크게 DOI 번호 획득 및 학술정보 관련 메타데이터 등록, 갱신, 유지, 해석과 같은 기본서비스, 그리고 Cited-by Linking 등과 같은 부가가치서비스로 구분된다. 기본서비스인 DOI 기탁은 CrossRef XML DTD를 기반으로 기탁 스키마를 이용하여 논문의 메타데이터를 생성하고, 이를 CrossRef에 제공하는 과정으로 수행된다. Cited-by Linking은 논문의 피인용 추적 서비스로 CrossRef에서 DOI번호가 부여된 논문 1건당 피인용 횟수의 통계 데이터를 제공하는 서비스이다.

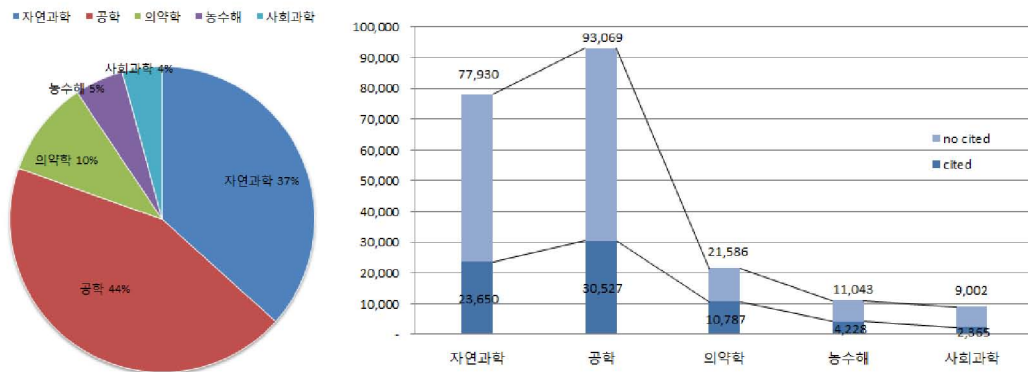
2016년 3월 기준 DOI 참여 학술지는 총 445종이고, DOI 기탁 논문은 약 21만 건이다. 2008년부터 2016년 3월까지 전체 기탁 현황에 관한 DOI 기탁 논문 건수와 누적 건수는 <표 1>과 같다. 학술지의 DOI 참여가 증가하면서 기탁 논문 누적 건수는 꾸준히 증가하고 있다.

2008년부터 본 연구의 수행시점인 2016년 3월까지 누적된 전체 DOI 기탁 논문의 주제별²⁾ 분포 현황은 <그림 1>과 같다. 전체 DOI 기탁 논문 약 21만 건 중 자연과학 분야 77,930건(142종), 공학 분야 93,069건(173종)으로 총 170,999건(315종)으로 약 80%를 구성하고 있는 것을 확인할 수 있다.

<표 1> DOI 기탁 현황

	2008	2009	2010	2011	2012	2013	2014	2015	2016
DOI 기탁 논문 건수	1,480	4,537	40,679	25,208	28,354	33,412	33,522	30,800	14,477
DOI 기탁 논문 누적 건수	1,480	6,017	46,696	71,904	100,419	133,831	167,353	198,153	212,630

2) 본 연구에서는 WoS의 세부주제를 적용하였으며, 이를 OECD에서 제시하는 대분류와 중분류에 매핑 하였다. OECD Classification- Web of Science Subject Headings (http://incites.isiknowledge.com/common/help/h_field_category_oecd_wos.html)



〈그림 1〉 DOI 기탁 논문의 주제별 분포 현황 및 1번 이상 피인용 된 논문 수

3.2 표본 데이터 수집

본 연구에서 활용하는 Cited-by Linking 서비스는 DOI 참여 학술지를 대상으로 CrossRef에서 제공하는 조건부, 무료 서비스이다. CrossRef에서는 Cited-by Linking 서비스 제공에 동의한 기관에게 기탁된 논문들을 인용한 논문들(Cited-by Linking)의 데이터를 제공하고 있고, KISTI에서는 월 단위로 학술정보통합관리시스템에 이 데이터를 구축, 관리하고 있다. 전체 기탁된 21만 건의 논문 중 Cited-by Linking 서비스를 통해 수집된 2016년 3월까지 누적된 적어도 1회 이상 피인용 된 논문 건수는 총 71,578건이며, 피인용 횟수는 총 197,014번이다.

앞서 살펴본 바와 같이 DOI 기탁 논문의 주제별 분포 현황에서 자연과학과 공학 분야가 큰 비중을 차지하고 있음을 확인하였다. 이에 본 연구에서는 자연과학과 공학 분야의 DOI 기탁 논문과 그 논문을 인용한 논문들을 분석 대상으로 한다.

분석 대상이 되는 자연과학과 공학분야에서 1번 이상 인용을 받은 논문 건수는 각각 23,650건

(피인용횟수 67,801번), 30,527건(피인용 횟수 71,708번)으로 총 54,177건(피인용 횟수 139,509번)이다. CrossRef에서 제공한 Cited-by Linking 서비스를 통해 수집된 raw data는 기탁된 메타데이터 형식을 그대로 활용한 자료이다. 그러므로 분석을 위하여, 인용 받은 논문의 출판년도보다 인용한 논문의 출판년도가 더 작은 값인 경우를 확인하였다. 이에 해당하는 논문 건수는 21건(자연과학 7건, 공학 14건), 피인용 횟수로는 61번으로 분석대상에서 제외하는 전처리 과정을 실시하였다. 이를 OECD 중분류별 표로 요약하면 〈표 2〉와 같다.

〈표 2〉를 바탕으로 분석 대상이 되는 자연과학, 공학 분야의 기탁 논문들과 이 중에서 1번 이상 인용된 논문들(이하 Source DOI)의 논문 단위 “사용 언어 구분”, 학술지 단위의 “SCIE와 SCOPUS 등재 여부” 데이터를 수집하였다. 그리고 Source DOI들을 1번 이상 인용한 해외 또는 국내 논문들(이하 Link DOI) 중 SCIE 논문의 주제 분야를 분석하기 위하여 KISTI의 e-Gate 시스템을 통해 “SCIE 등재 여부”를 식별하였다.

〈표 2〉 자연과학 및 공학 분야 DOI 논문 기탁, Cited-by Linking 현황

OECD 분류		학술지 종수	기탁 논문수	기탁 논문 비율	1번 이상 인용된 논문수	1번 이상 인용된 논문 비율	피인용 횟수	피인용 비율
자연 과학	수학	18	8,453	10.8%	2,645	11.2%	6,893	10.2%
	컴퓨터 공학	25	12,313	15.8%	3,065	13%	6,787	10%
	물리학, 천문학	13	5,723	7.3%	2,028	8.6%	5,395	8%
	화학	15	10,859	13.9%	1,003	4.2%	2,585	3.8%
	지구과학, 환경	28	10,925	14%	4,528	19.2%	10,751	15.9%
	생물학	28	14,315	18.4%	6,510	27.5%	28,648	42.3%
	기타	15	15,342	19.7%	3,864	16.3%	6,712	9.9%
소 계		142	77,930	100%	23,643	100%	67,771	100%
공학	토목 공학	34	13,429	14.4%	5,388	17.7%	19,048	26.6%
	전기, 전자 공학	32	25,656	27.6%	6,388	20.9%	12,878	18%
	기계 공학	24	15,344	16.5%	4,101	13.4%	7,283	10.2%
	화학 공학	4	2,581	2.8%	1,031	3.4%	2,482	3.5%
	재료 공학	21	13,141	14.1%	5,083	16.7%	11,015	15.4%
	의료 공학	2	229	0.2%	34	0.1%	44	0.1%
	환경 공학	26	10,648	11.4%	3,530	11.6%	6,957	9.7%
	환경 생명 공학	2	838	0.9%	370	1.2%	1,050	1.5%
	산업 생명 공학	1	126	0.1%	37	0.1%	54	0.1%
	나노 과학	1	64	0.1%	21	0.1%	76	0.1%
소 계		173	93,069	100%	30,513	100%	71,677	100%
총 합계		315	170,999		54,156		139,448	

Source DOI 총 54,156건은 Link DOI 총 139,448건에 의해 인용을 받았다. Link DOI 주제 분야를 확인하기 위하여, 총 139,448건 중 “SCIE 등재 여부”를 식별한 결과, e-Gate를 통해 식별되지 않은 비율은 전체 15%였으며, 85%는 SCIE 해당여부의 식별이 가능하였다. 식별된 85%를 100으로 기준할 때, SCIE에 해당하는 논문은 40%, SCIE에 해당하지 않는 논문은 60%였다. 이 40%는 자연과학 분야에 해당하는 Source DOI를 인용한 Link DOI 중 SCIE급 논문 24,717건, 공학 분야에 해당하는 Source DOI를 인용한 Link DOI 중 SCIE급 논문 22,981건으로 구성되어있다. 이와 같은 Link DOI 중 SCIE 급 총 47,698건을 대상으로

주제적 연관성 분석을 실시하기 위해, 47,698건이 속하는 학술지명을 확인하여, JCR 2014 기준으로 제시된 주제 분야를 적용하였다. JCR에서 주제 분야가 복수로 제시된 경우에는 첫 번째 제시된 주제 분야를 사용하였다.

4. Cited-by Linking 데이터 분석 결과

4.1 인지적 접근성 (Cognitive accessibility) - 사용 언어 구성비

본 절에서는 분석 데이터를 기반으로 인용

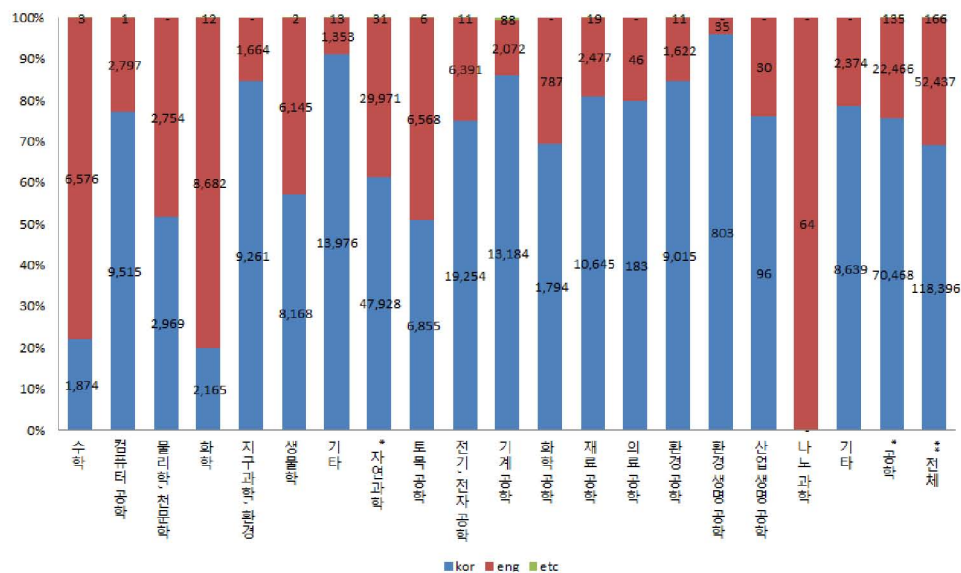
의 다양한 요인 중 논문의 언어를 이해 할 수 있을 때, 즉 논문을 인지하는 것에 대한 어려움이 없을 때, 더 많이 인용될 것이라는 가정을 검증하도록 한다. 먼저, 자연과학과 공학 분야의 전체 기탁한 논문의 사용 언어 비율은 다음 <그림 2>와 같다.

자연과학 분야의 기탁한 논문 77,930건 중 61.5%(47,928건)가 국문, 38.46%(29,971건)가 영문, 그리고 0.04%가 기타 언어로 구성되어 있다. 공학 분야의 경우는 기탁한 논문 93,069건 중 75.72%(70,468건)의 국문, 24.14%(22,466건)의 영문, 그리고 0.15%(135건)가 기타 언어로 구성되어 있다. 전반적으로 공학 분야 보다는 자연과학 분야의 논문이 영문으로 작성된 것이 더 많았으며, 자연과학 분야에서도 수학의 경우 영문 논문이 더 많이 작성된 것을 확인할 수 있다. 물리학·천문학의 경우에는 국문이 51.88%, 영문이 48.12%로 거의 비등한 구

성비를 보이고 있다. 공학 분야 중, 토목 공학의 경우에 국문 51.05%, 영문 48.91%로 비등한 구성비를 보이고 있으며, 나노 과학 분야의 경우, 이 분야에 해당하는 2종(Nano-성균관대학교 성균나노과학기술원, Advances in nano research-테크노프레스) 모두 영문지로 구성되어 있음을 확인하였다.

DOI 참여 학술지 중 1번이라도 인용 받은 논문들의 사용 언어 비율과 DOI 참여 학술지 중 인용 받지 않은 논문들의 사용 언어 비율을 분석하면 <표 3>과 같다. 직관적으로 확인 가능하도록 각각 구성 비율을 표로 제시하였으며, <그림 3>과 같이 그래프로 표현하였다.

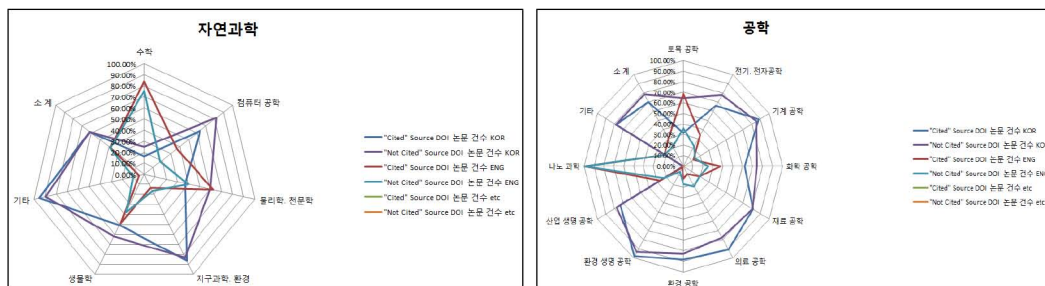
자연과학 분야 중 수학에 해당하는 논문의 경우, 1번 이상 인용 받은 논문은 <표 2>에서 확인하는 것과 같이 2,645건이다. 이 중 <표 3>과 같이 국문은 436건으로 16.48%를 차지했고, 영문은 2,208건으로 83.48%를 차지했다. 전체



<그림 2> 자연과학과 공학 분야의 전체 기탁 논문 사용 언어 비율

〈표 3〉 자연과학과 공학 분야의 논문 사용 언어 구성 비율

OECD주제분류		Cited Source DOI 논문 건수 비율			Not Cited Source DOI 논문 건수 비율			Cited Source DOI 피인용수		
		KOR	ENG	etc	KOR	ENG	etc	KOR	ENG	etc
자연과학	수학	16.48%	83.48%	0.04%	24.78%	75.19%	0.03%	741	6,151	1
	컴퓨터 공학	63.03%	36.97%	-	82.01%	17.98%	0.01%	3,256	3,531	-
	물리학·천문학 화학 ³⁾	37.08%	62.92%	-	60%	40%	-	1,312	4,083	-
	지구과학·환경	86.48%	13.52%	-	83.55%	16.45%	-	9,000	1,751	-
	생물학	50.77%	49.23%	-	62.31%	37.67%	0.03%	8,783	19,865	-
	기타	95.24%	4.76%	-	89.70%	10.18%	0.11%	6,398	314	-
	소 계	61.93%	38.07%	-	71.45%	28.51%	0.04%	29,490	35,695	1
공학	토목 공학	31.24%	68.76%	-	64.31%	35.61%	0.07%	2,949	16,099	-
	전기·전자공학	66.08%	33.92%	-	78.02%	21.92%	0.06%	6,919	5,959	-
	기계 공학	88.12%	11.88%	-	85.12%	14.10%	0.78%	6,091	1,192	-
	화학 공학	62.27%	37.73%	-	74.27%	25.73%	-	1,317	1,165	-
	재료 공학	81.09%	18.91%	-	80.96%	18.81%	0.24%	8,663	2,352	-
	의료 공학	91.18%	8.82%	-	77.95%	22.05%	-	36	8	-
	환경 공학	88.19%	11.81%	-	82.92%	16.93%	0.15%	6,126	831	-
	환경 생명 공학	98.38%	1.62%	-	93.80%	6.20%	-	1,037	13	-
	산업 생명 공학	72.97%	27.03%	-	77.53%	22.47%	-	43	11	-
	기타	78.48%	21.52%	-	78.42%	21.58%	-	8,310	2,480	-
	소 계	70.04%	29.96%	-	78.48%	21.30%	0.22%	41,491	30,186	-
합 계		66.588%	33.412%	0.002%	75.56%	24.30%	0.14%	70,981	65,881	1



〈그림 3〉 자연과학과 공학 분야의 논문 사용 언어 구성 비율 그래프

기탁된 논문 중 1번 이상 인용 받은 논문 건수를 제외하면, 인용 받지 않은 논문 건수를 확인할 수 있다. 인용 받지 않은 논문들의 사용 언어

구성 비율 역시 〈표 3〉에서 확인 가능하다.

논문의 언어를 이해할 수 있을 때, 더 많이 인용될 것이라는 가정에 관한 검증은 위하여

3) 데이터 미확보로 인한 분석 대상 제외

위와 같은 1번 이상 인용 받은 논문 건수와 인용 받지 않은 논문 건수의 주제별 사용 언어 구성비 데이터를 비교한 결과 자연과학과 공학 분야의 서로 비슷한 양상을 확인할 수 있다. 자연과학의 경우, 6개의 하위 주제 분야 중 4개의 하위 주제 분야(수학, 컴퓨터공학, 물리학·천문학, 생물학)에서 영문으로 작성된 논문의 인용 비율이 더 높게 나타났으며, 공학의 경우에는 10개의 하위 주제 분야 중 5개(토목 공학, 전기·전자공학, 화학공학, 재료공학, 산업생명공학)가 영문으로 작성된 논문의 인용 비율이 더 높게 나타났다. 나노 과학 분야의 경우에는 영문으로 작성된 논문만이 기탁되었으므로 비교 대상에서 제외하였다.

화학 분야에 해당하는 학술데이터가 해외 출판사로 이관되면서 이관 시점부터 CrossRef를 통해 Cited-by Linking 데이터를 확보되지 못하였으므로 화학 분야는 분석대상에서 제외하였다. 화학 분야를 제외하고, 자연과학 분야 소계를 비교하였을 때, Cited Source DOI의 국문: 영문의 비율이 61.93: 38.07, Not Cited Source DOI의 국문: 영문은 71.45: 28.51 이었다. 자연과학 분야에서 영문으로 작성된 논문의 인용 비율이 9.56%(38.07-28.51) 공학 분야의 소계에서는 8.38%(29.96-21.30) 차이를 보이며, 두 분야 모두 영문으로 작성된 논문의 인용 비율이 더 높음을 확인하였다.

4.2 품질 및 중요성(Perceived quality & Significance) – SCIE, SCOPUS 등재 여부 구성비

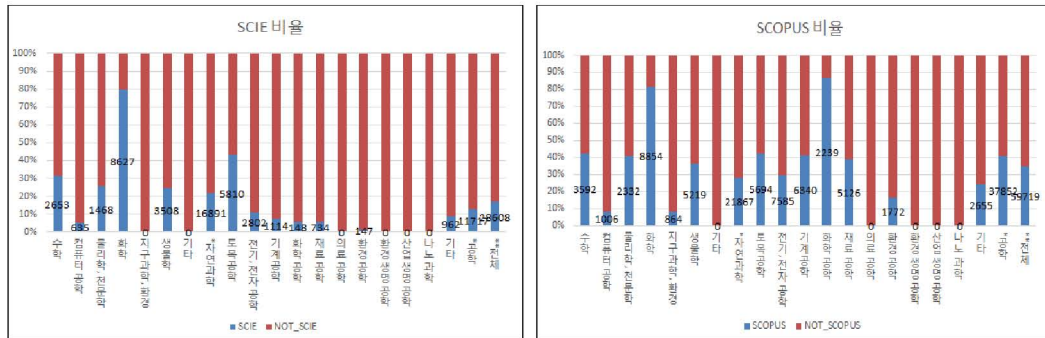
품질 및 중요성의 측면은 논문의 수준이 높

거나 과학적으로 더 중요한 경우, 논문의 인용이 더 높아질 것으로 접근한다. 이와 같은 접근을 검증하기 위해서 SCIE와 SCOPUS 등재지의 경우 더 많이 인용될 것이라는 가설을 제시하였고, 이를 위해 본 절에서는 DOI 참여 학술지 중 1번이라도 인용 받은 논문들의 SCIE와 SCOPUS 등재 여부 비율과 DOI 참여 학술지 중 인용 받지 않은 논문들의 SCIE와 SCOPUS 등재 여부 비율을 분석한다.

먼저, 자연과학과 공학 분야의 전체 기탁한 논문의 SCIE, SCOPUS 해당 여부 비율은 다음 〈그림 4〉와 같다. 자연과학 분야에서는 기탁한 논문 77,930건 중 21.67%(16,891건)가 SCIE에 등재, 28.06%(21,867건)가 SCOPUS에 등재되어 있다. 공학 분야의 경우는 기탁한 논문 93,069건 중 12.59%(11,717건)가 SCIE 등재, 40.67%(37,852건)가 SCOPUS에 등재되어 있다. 전반적으로 자연과학, 공학 분야 모두 SCIE보다 SCOPUS에 등재된 DOI 참여 학술지가 더 많은 것을 확인할 수 있다.

〈표 4〉는 SCIE 등재 여부를 Cited Source DOI 논문 건수와 Not Cited Source DOI 논문 건수의 비율, 그리고 인용된 경우 Cited Source DOI의 피인용수를 제시한 것이다. 앞 절에서 언급한 것과 같이 자연과학 분야 중 화학은 분석대상에서 제외한다. 또한 자연과학 분야 중 지구과학·환경, 기타 그리고 공학 분야 중 의료 공학, 환경 생명 공학, 산업 생명 공학, 나노 과학의 경우에도 기탁된 SCIE 논문이 없으므로 분석 대상에서 제외한다.

1번 이상 인용 받은 논문 건수와 인용 받지 않은 논문 건수의 주제별 SCIE 등재 여부 구성비를 비교한 결과, 자연과학의 경우에 SCIE 등



〈그림 4〉 자연과학과 공학 분야의 전체 기탁 논문 SCIE, SCOPUS 논문 비율

〈표 4〉 자연과학과 공학 분야의 SCIE 등재 여부 구성 비율

OECD주제분류		Cited Source DOI SCIE급 논문 건수 비율		Not Cited Source DOI SCIE급 논문 건수 비율		Cited Source DOI SCIE급 피인용수	
		N	Y	N	Y	N	Y
자연과학	수학	52.36%	47.64%	76.05%	23.95%	2,679	4,214
	컴퓨터 공학	87.11%	12.89%	97.40%	2.60%	5,450	1,337
	물리학·천문학	55.57%	44.43%	84.65%	15.35%	2,045	3,350
	생물학	63.24%	36.76%	85.71%	14.29%	9,846	18,802
소 계		65.27%	34.73%	87.53%	12.47%	20,020	27,703
공학	토목 공학	33.87%	66.13%	72.05%	27.95%	3,199	15,849
	전기·전자공학	78.07%	21.93%	92.73%	7.27%	8,672	4,206
	기계 공학	96.63%	3.37%	91.32%	8.68%	7,029	254
	화학 공학	93.50%	6.50%	94.76%	5.24%	2,296	186
	재료 공학	92.78%	7.22%	95.46%	4.54%	10,072	943
	환경 공학	100%	0%	97.93%	2.07%	6,957	0
	기타	87.99%	12.01%	93.55%	6.45%	9,094	1,696
소 계		79.77%	20.23%	90.87%	9.13%	47,319	23,134
합 계		75.10%	24.90%	89.87%	10.13%	67,339	50,837

재지가 있는 4개의 하위 주제 분야(수학, 컴퓨터공학, 물리학·천문학, 생물학) 모두 SCIE 등재 논문의 인용 비율이 더 높게 나타났다. 공학의 경우에는 SCIE 등재지가 있는 7개의 하위 주제 분야 중 5개(토목 공학, 전기·전자공학, 화학공학, 재료공학, 기타)가 SCIE 등재 논

문의 인용 비율이 더 높게 나타났다. 기계 공학과 환경 공학의 경우에는 SCIE 등재 논문이지만 1번 이상 인용되지 않은 논문 건수 비율이 근소하게 더 높았다.

SCIE급 논문의 피인용수와 SCIE가 아닌 논문의 피인용 수를 비교하였을 때, SCIE급 논문

의 피인용수가 월등하게 높을 것을 예상하였다. 그러나 자연과학 분야의 경우에는 세부 분야별 차이를 보이며 컴퓨터 공학을 제외하고 모두 높았지만, 공학의 경우에는 토목 공학만이 SCIE 급 논문의 피인용 수가 월등하게 높았다. 공학의 경우에는 1번 이상 인용 되는 논문의 SCIE 급 비율이 1번도 인용되지 않은 논문의 SCIE 급 비율보다 5개 분야에서 높았음에도 불구하고, 피인용수에서는 오히려 SCIE급 논문보다 SCIE급이 아닌 논문들의 피인용수가 현저하게 높았다.

〈표 5〉는 SCOPUS 등재 여부를 Cited Source DOI 논문 건수와 Not Cited Source DOI 논문 건수의 비율, 그리고 인용된 경우 Cited Source DOI의 피인용수를 제시한 것이다. SCOPUS의 경우에는 자연과학 분야 중 기타, 공학 분야 중

의료 공학, 환경 생명 공학, 산업 생명 공학, 나노 과학 분야에 해당하는 DOI 참여 학술지 중 SCOPUS에 등재된 학술지가 없음을 〈그림 4〉에서 확인 가능하다. 그러므로 이 분야들은 분석 대상에서 제외하도록 한다.

1번 이상 인용 받은 논문 건수와 인용 받지 않은 논문 건수의 주제별 SCOPUS 등재 여부 구성비를 비교한 결과, 자연과학의 경우에 SCOPUS 등재지가 있는 5개의 하위 주제 분야(수학, 컴퓨터공학, 물리학·천문학, 지구과학·환경, 생물학) 모두 SCOPUS 등재 논문의 인용 비율이 더 높게 나타났다. 공학의 경우에는 SCOPUS 등재지가 있는 7개의 하위 주제 분야 중 5개(토목 공학, 전기·전자공학, 화학 공학, 환경 공학, 기타)가 SCOPUS 등재 논문의 인용 비율이 더 높게 나타났다. 기계 공학과 재료 공학의 경우에

〈표 5〉 자연과학과 공학 분야의 SCOPUS 등재 여부 구성 비율

OECD주제분류		Cited Source DOI SCOPUS급 논문 건수 비율		Not Cited Source DOI SCOPUS급 논문 건수 비율		Cited Source DOI SCOPUS급 피인용수	
		N	Y	N	Y	N	Y
자연과학	수학	40.11%	59.89%	65.47%	34.53%	1,933	4,960
	컴퓨터 공학	79.15%	20.85%	96.03%	3.97%	4,425	2,362
	물리학·천문학	38.86%	61.14%	70.45%	29.55%	1,363	4,032
	지구과학·환경	91.48%	8.52%	92.53%	7.47%	9,570	1,181
	생물학	51.21%	48.79%	73.82%	26.18%	7,994	20,654
소 계		62.59%	37.41%	81.84%	18.16%	25,285	33,189
공학	토목 공학	33.87%	66.13%	73.49%	26.51%	3,199	15,849
	전기·전자공학	55.89%	44.11%	75.26%	24.74%	5,764	7,114
	기계 공학	60.47%	39.53%	58.02%	41.98%	4,484	2,799
	화학 공학	6.89%	93.11%	17.39%	82.61%	97	2,385
	재료 공학	61.66%	38.34%	60.59%	39.41%	6,663	4,352
	환경 공학	74.73%	25.27%	88.37%	11.63%	4,570	2,387
	기타	67.88%	32.12%	81.48%	18.52%	6,536	4,434
소 계		55.88%	44.12%	70.82%	29.18%	31,133	39,320
합 계		52.71%	37.45%	74.63%	25.37%	56,598	72,509

는 SCOPUS 등재 논문이지만 인용되지 않은 논문 건수 비율이 근소하게 더 높았다. SCOPUS급 논문의 피인용수와 SCOPUS급이 아닌 논문의 피인용 수를 비교하였을 때는 자연과학에서 생물학 분야, 공학에서 토목 공학 분야의 SCOPUS급 논문의 피인용수가 두드러지게 높은 것을 확인할 수 있었으며, 각 세부주제 분야별 차이가 있었다.

4.3 주제 분야 유사성 - SCIE 주제 분야 적용

본 절에서는 인용 요인 중, 주제적 연관성을 확인하기 위한 측면으로 다수의 논문들이 유사한 주제 분야에서 인용 되었을 것을 검증한다. 이를 위하여 DOI 참여 학술지에 속한 논문들(Source DOI)을 인용한 해외 또는 국내 논문들(Link DOI) 중 SCIE 논문의 주제 분야를 분석한다.

앞선 3.2와 같이 수집된 분석 대상이 되는 Link DOI 중 SCIE급 총 47,698건의 논문 중, 46,898건의 논문의 JCR 주제 분야를 식별하였다. 본 연구에서 해외 또는 국내 논문들 중 SCIE 논문을 분석 대상으로 정한 것은 WoS에서 제공하는 주제 분야를 적용하여 보다 객관적인 결과를 얻기 위함이며, 이를 앞서 언급된 OECD 주제 분야와 매핑하였다. 자연과학 6개의 세부 주제 분야(화학 분야 제외)에 해당하는 국내 Source DOI를 인용한 Link DOI와 공학 11개의 세부주제 분야에 해당하는 국내 Source DOI를 인용한 Link DOI가 속하는 OECD 주제 분야는 총 29개였다.

이를 정리하여 <표 6>과 같이 옆에 Source

DOI(인용 받은)를, 행에 Link DOI(인용한)를 표시하였다. <표 6>에서 확인할 수 있듯이 대부분 유사한 분야에서 Source DOI를 인용하였지만, 각 세부주제 분야별로 특징이 있었다. <표 6>에 피인용수가 가장 큰 값을 표시하였으며, 이를 통해 자연과학 분야와 공학 분야의 차이를 확인할 수 있었다. 자연과학 분야의 경우에는 대부분 최댓값이 대각선으로 나타나며, 이는 거의 동일한 분야에서 인용된 것으로 볼 수 있다. 그러나 공학 분야의 경우, 총 11개의 세부 분야 중 5개 분야(토목 공학, 전기·전자 공학, 기계 공학, 재료 공학, 기타 공학 분야)에서 동일 분야의 피인용 최댓값을 가졌고, 6개는 타분야에서 피인용 최댓값을 가진 것을 확인할 수 있다.

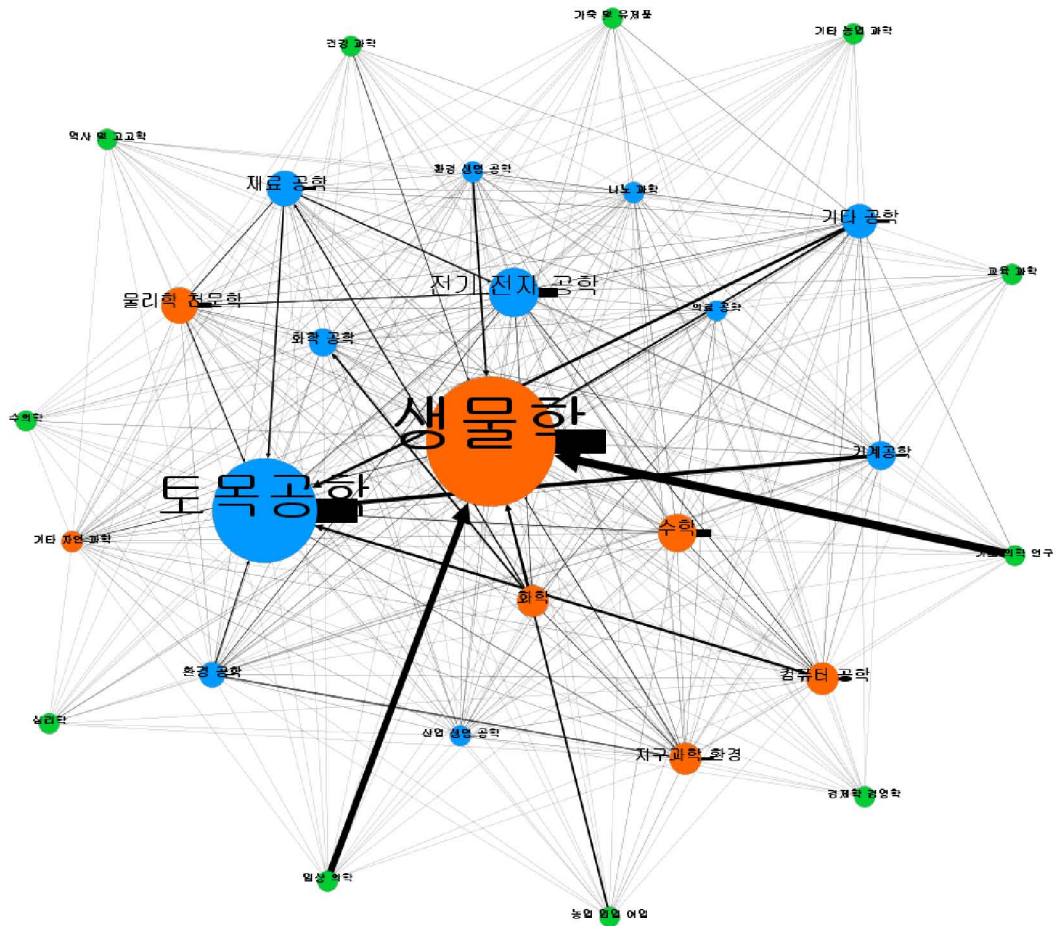
다음은 각 주제 간의 네트워크를 통하여 그 관계 파악을 용이하게 하기 위해 <그림 5>와 같이 <표 6>을 매트릭스로 하여 Gephi를 활용하여 시각화를 하였다. 이 네트워크는 인용 관계 네트워크로 피인용 연결(In-degree)에 의해 Node의 크기를 표시하고, 자연과학 분야는 오렌지색, 공학 분야는 하늘색으로 표시하였다. 연두색은 자연과학이나 공학이 아닌 타분야를 표시한다. 노드의 크기가 클수록 피인용수가 많은 주제이다. 자연과학 분야 중 생물학, 공학 분야 중 토목 공학의 경우, 동일한 주제(노드 옆의 굵은 ■ 표시)에서 피인용수를 많이 받았으며, 이 두 하위 분야의 경우 실제 데이터가 제기된 가설을 객관적으로 증명하고 있다. 또한 네트워크 지도를 통해서 SCIE급 해외 또는 국내 논문들이 국내의 생물학, 토목공학, 전기·전자 공학 분야의 논문을 많이 인용하고 있음을 한 눈에 확인할 수 있다.

〈표 6〉 자연과학과 공학 분야의 SCIE Source DOI의 WoS 주제별 피인용수

	자연 과학 Source DOI						공학 Source DOI										
	1.01	1.02	1.03	1.05	1.06	1.07	2.01	2.02	2.03	2.04	2.05	2.06	2.07	2.08	2.09	2.10	2.11
1.01	2084	55	1			1	207	9	25		4						2
1.02	133	661	9	7	1	31	755	86	31		39		13				32
1.03	9	5	1593	16	23	4	328	248	85	23	248	1	11			2	5
1.04	1	18	73	205	610	3	80	206	13	572	416	3	26	8	1	7	157
1.05	3	5	20	606	219	1	89	5		12	4		173				3
1.06	16	12	24	105	7033	2	12	9	5	50	5		98	21		2	148
1.07	4	4	12	3	204		16	10	5		4		2	1			1
2.01	3	390	6	88	1	3	7030	16	110	2	122		60		1	10	12
2.02	98	188	251	1		9	139	2636	57	3	159	2	8	1		3	9
2.03	9	71	31	20	5	4	1215	53	515	15	42		67			2	23
2.04		5	2	46	10	1	9	8	18	117	17		4			2	12
2.05		75	63	16	15	6	477	348	76	36	862		12		1	1	10
2.06	1	5	3		144	1	23	14	6		12	2					4
2.07	1	27	11	197	28	1	332	187	57	33	25		122				13
2.08	4	4	2	617		2	2	2	4	3		7	20			46	755
2.09				2		1				3							6
2.10		1	17	6	26		12	29	4	2	27					2	
2.11	28	57	48	27	548	7	868	116	143	46	59		28	5	2	8	1056
3.01	2	2	2	11	2499		1			58	1		3	3			101
3.02	2	6	9	2	2033	7	2	1	1		8	1	2	2			28
3.03	3	7	1	4	230	40	1			3	5						73
4.01	1	1	3	71	594		1	3	1		4		13	12			49
4.02	2		1	3	75								1	1			102
4.03					22								1	1			1
4.05			1	48					7	5		5	5			6	118
5.01					71			1									3
5.02	16	2					3	2									8
5.03	2					2	1	1									
6.01																	1

1.01	1.02	1.03	1.05	1.06	1.07	2.01	2.02	2.03	2.04	2.05	2.06	2.07	2.08	2.09	2.10	2.11
수학	컴퓨터 공학	물리학 천문학	지구 과학 환경	생물학	기타 자연 과학	토목 공학	전기 전자 공학	기계 공학	화학 공학	재료 공학	의료 공학	환경 공학	환경 생명 공학	산업 생명 공학	나노 과학	기타 공학

3.01	3.02	3.03	4.01	4.02	4.03	4.05	5.01	5.02	5.03	6.01
기초 의학 연구	임상 의학	건강 과학	농업 임업 어업	가축 및 유제품	수의학	기타 농업 과학	심리학	경제학 경영학	교육 과학	역사 및 고고학



〈그림 5〉 29개 주제 분야의 인용 네트워크 지도

5. 결론

인용은 복합적인 요인에 의한 지적인 관례이며, 본 연구에서는 자연과학과 공학 분야의 인용 패턴을 확인하기 위하여, Cited-by Linking 데이터를 활용하였다. Cited-by Linking 서비스는 CrossRef에서 제공하는 주요 서비스 중 하나로 해당 논문이 DOI를 통하여 얼마나 인용되었는지 추적하여 집계된 데이터이다. 이 범

위는 WoS, SCOPUS 그리고 국내 학술데이터베이스인 KSCD 보다 더 광범위하다. 그러므로 계량정보학 분야에서 인용 분석 대상으로 보다 객관적인 결과의 도출이 가능할 것으로 판단하였다. 이에 KISTI의 학술정보통합관리시스템에서 월단위로 구축·관리하는 Cited-by Linking 데이터를 분석대상으로 하여, 인지적 접근성의 측면과 품질 및 중요성의 측면에서 제시된 가설과 마지막으로 유사한 주제 분야에서 인용이

일어날 것이라는 가설을 제시하여 검증을 수행하였다. 분석 대상의 범위는 전체 기탁된 21만 건의 논문 중 자연과학 분야 77,930건(142종), 공학 분야 93,069건(173종)으로 총 170,999건(315종)이며, 2016년 3월까지 Cited-by Linking 서비스를 통해 수집된 Link DOI이다.

각 측면의 가설은 다음과 같으며, 자연과학과 공학 분야의 OECD 세부주제 분야 별로 가설의 결과를 제시하였다.

(1) 인지적 접근성 - 영문지가 국문지보다 더 인용이 많이 될 것이다.

이와 같은 가설을 이를 검증하기 위해 DOI 참여 학술지 중 1번이라도 인용 받은 논문들의 사용 언어 비율과 DOI 참여 학술지 중 인용 받지 않은 논문들의 사용 언어 비율을 분석하였다. 그 결과, 자연과학의 경우 6개의 하위 주제 분야 중 4개의 하위 주제 분야에서 국문:영문 비가 “1번이라도 인용 받은 Source DOI”가 “인용 받지 않은 Source DOI”보다 영문(논문건수)이 차지하는 비율이 더 높게 나타났다. 공학의 경우에는 11개의 하위 주제 분야 중 5개의 하위 주제 분야에서 위와 같은 결과를 얻었다.

이 가설은 Cite-by Linking 데이터가 글로벌하게 인용되며, 그 횟수를 누적적으로 집계하는 것을 착안하여, 자연과학과 공학 분야의 모든 하위 주제 분야에서 “1번이라도 인용 받은 Source DOI”이 국문:영문에서 영문(논문건수)의 비율이 높을 것으로 예상하고 세운 것이다. 자연과학과 공학분야 모두 영문으로 작성된 논문의 인용 비율이 각각 더 높음을 확인하였다.

(2) 품질 및 중요성 - SCIE와 SCOPUS 등재지의 경우 더 많이 인용 될 것이다. 이를 검증하기 위해 DOI 참여 학술지 중 1번이라도 인용 받은 논문들의 SCIE와 SCOPUS 등재 여부 비율과 DOI 참여 학술지 중 인용 받지 않은 논문들의 SCIE와 SCOPUS 등재 여부 비율을 분석하였다. 각각 주제별 SCIE 등재 여부 구성비를 비교한 결과, 자연과학의 경우에 SCIE 등재지가 있는 4개의 하위 주제 분야 모두 SCIE 등재 논문의 인용 비율이 더 높게 나타났다. 공학의 경우에는 SCIE 등재지가 있는 7개의 하위 주제 분야 중 5개가 SCIE 등재 논문의 인용 비율이 더 높게 나타났다.

또한 “1번이라도 인용 받은 Source DOI” 중, SCIE급 논문의 피인용수와 SCIE가 아닌 논문의 피인용 수를 비교하였을 때, SCIE급 논문의 피인용수가 월등하게 높을 것을 예상하였다. 자연과학 분야의 경우에는 세부주제 분야별 차이를 보이며 컴퓨터 공학을 제외하고 모두 SCIE급 피인용수가 높았지만, 공학의 경우에는 토목 공학만이 SCIE급 논문의 피인용수가 월등하게 높을 뿐이었다. 공학의 경우에는 피인용수에서는 오히려 SCIE급 논문보다 SCIE급이 아닌 논문들의 피인용수가 현저하게 높았다.

각각 주제별 SCOPUS 등재 여부 구성비를 비교한 결과, 자연과학의 경우에 SCOPUS 등재지가 있는 5개의 하위 주제 분야 모두 SCOPUS 등재 논문의 인용 비율이 더 높게 나타났다. 공학의 경우에는 SCOPUS 등재지가 있는 7개의 하위 주제 분야 중 5개에서 SCOPUS 등재 논문의 인용 비율이 더 높게 나타났다. 그리고 “1번이라도 인용 받은 Source DOI” 중 SCOPUS급 논문의 피인용수와 SCOPUS급이 아닌 논문의 피인

용 수를 비교하였을 때는 자연과학 분야와 공학 분야 모두 세부주제 분야별 차이를 보였다.

(3) 주제적 연관성 - 다수의 논문들이 유사한 주제 분야에서 인용되었을 것이다. 이는 검증하기 위해 "Source DOI"를 인용한 해외 또는 국내 논문(Link DOI)들 중 SCIE급 논문을 표본으로 하여, 그 주제 분야를 분석하였다. Link DOI 중 SCIE급 총 47,698건의 논문 중, 46,898건의 논문의 JCR 주제 분야가 식별 가능하였다. 분석 결과, 자연과학 6개 세부주제 분야에 해당하는 "국내 Source DOI"를 인용한 "Link DOI"와 공학 분야 11개 세부주제 분야에 해당하는 "국내 Source DOI"를 인용한 "Link DOI"가 속한 OECD 주제 분야는 총 29개였다. 대부분 유사한 분야에서 Source DOI를 인용하였지만, 자연과학 분야와 공학 분야의 차이를 확인할 수 있었다. 자연과학 분야의 경우에는 거의 동일한 분야에서 인용되었으며, 공학 분야의 경우에는 11개의 세부주제 분야별로 차이를 보였다. 이는 자연과학보다 공학 분야에서 융합 연구가 더 많이 수행되는 학문적 특성이 반영된 것으로 추정된다.

본 연구에서는 위와 같이 다양한 인용 요인에 관하여 데이터를 기반으로 제기된 가설을 검증하기 위해 다각적인 분석을 수행하였다. Cited-by Linking 데이터를 활용하는 경우, 그 커버리지가 국내뿐만 아니라 해외 데이터베이스를 포함하여 더 광범위할 것으로 예상하고 이 연구를 시작하였다. 분석 과정 중 발견된 이슈사항으로 Cited-by Linking 데이터는 DOI 기탁을 하는 대형 기관에게 반출되기 때문에 DOI 참여 학회가 해외 출판사로 이관되는 경우, 중간에 데이터를 확보할 수 없다는 것을 파악할 수 있었다. 그러므로 이와 같은 이슈사항 등을 raw data의 전처리 과정에서 미리 확인해야 하는 제한점이 있다.

그러나 본 연구는 글로벌하게 수집된 데이터를 기반으로 자연과학과 공학 분야의 세부주제 분야별 인용 패턴을 파악하고, 인용 요인에 관한 객관적인 검증 데이터를 제공했다는 점에서 의의를 가진다. 향후, 자연과학과 공학 분야의 연구자들에게 자신들이 연구 성과를 더 널리 공유할 수 있게 하는 방향을 제시할 뿐만 아니라 인용 행태에 관한 연구에서도 유용하게 활용될 것으로 기대한다.

참 고 문 헌

- 김미라, 이응봉 (2013). 참고문헌 분석을 통한 국방과학 연구자들의 인용행태 연구. 한국문헌정보학회지, 47(2), 143-159. <http://dx.doi.org/10.4275/kslis.2013.47.2.143>
- 김병규, 강무영, 최선희, 김수영, 유범중, 신재도 (2011). KSCD를 활용한 국내 과학기술자의 해외 학술지 인용행태 연구. 정보관리학회지, 28(2), 117-133. <http://dx.doi.org/10.3743/kosim.2011.28.2.117>

- 서태설, 최희운 (2011). DOI와 오픈엑세스를 활용한 학술지의 국제적 이용 활성화 방안. *정보관리연구*, 42(4), 1-21. <http://dx.doi.org/10.1633/jim.2011.42.4.001>
- 장덕현, 장환석 (2007). 기계공학 연구자들의 인용행태 분석: P대학 기계공학부 박사학위 논문을 중심으로. *정보관리연구*, 38(3), 111-135. <http://dx.doi.org/10.1633/jim.2007.38.3.111>
- 정은경, 김병규, 전지영, 최선희, 강무영 (2011). CrossRef DOI 기반 Cited-by linking 결과의 학술정보 생산자 활용시스템 연구. *한국정보처리학회*, 18(1), 1573-1575.
- 조현양, 조현선 (2005). 주요 4개 공학분야 연구자의 문헌인용 행태 연구. *정보관리연구*, 36(2), 1-24. <http://dx.doi.org/10.1633/jim.2005.36.2.001>
- Cole, S. (1992). *Making science: between nature and society*. Cambridge, MA: Harvard University Press.
- Davis, P. M., & Cohen, S. A. (2001). The effect of the Web on undergraduate citation behavior 1996-1999. *Journal of the American Society for Information Science and Technology*, 52(4), 309-314.
- Gross, P. L. K., & Gross, E. M. (1927). College libraries and chemical education. *Science*, 66(1713), 385-389. <http://dx.doi.org/10.1126/science.66.1713.385>
- Liu, Z. (1997). Citation theories in the framework of international flow of information: New evidence with translation analysis. *Journal of the American Society of Information Science*, 48(1), 80-87.
- Lutz, B., & Hans-deiter, D. (2008). What do citation counts measure? A review of studies on citing behavior. *Journal of Documentation*, 64(1), 45-80. <http://dx.doi.org/10.1108/00220410810844150>
- Milojević, S. (2012). How are academic age, productivity and collaboration related to citing behavior of researchers?. *PloS one*, 7(11), e49176. <http://dx.doi.org/10.1371/journal.pone.0049176>
- Obuh, A. O., & Babatope, I. S. (2011). Student citation behaviour in Delta State University, Abraka, Nigeria. *Library Philosophy and Practice*. Retrieved from <http://unllib.unl.edu/LPP/obuh-babatope.htm>
- Seo, Tae-Sul, Jung, Eun-Gyeong, & Kim, Hwanmin (2013). Patterns of citing Korean DOI journals according to CrossRef's Cited-by Linking and a local journal citation database. *Journal of Information Science Theory and Practice*, 1(2), 58-68. <http://dx.doi.org/10.1633/JISTaP.2013.1.2.4>
- Van Raan, A. F. J. (2005). Fatal attraction: Conceptual and methodological problems in the ranking of universities by bibliometric methods. *Scientometrics*, 62, 133-143. <http://dx.doi.org/10.1007/s11192-005-0008-6>

Geghi. [Computer software]. Retrieved from <http://gephi.github.io/>

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

- Chang, Duk-Hyun, & Jang, Hwan-Seok (2007). A research on citing behaviors of researchers in mechanical engineering. *Journal of Information Management*, 38(3), 111-135.
<http://dx.doi.org/10.1633/jim.2007.38.3.111>
- Cho, Hyun-Yang, & Cho, Hyun-Sun (2005). A comparative study on the citing behavior of scholars in four major engineering fields. *Journal of Information Management*, 36(2), 1-24.
<http://dx.doi.org/10.1633/jim.2005.36.2.001>
- Jung, Eun-Kyoung, Kim, Byung-Kyu, Jeon, Ji-Young, Choi, Seon-Heui, & Kang, Mu-Yeong (2011). Application system plan for scholarly information producer using the result of cited-by linking based on CrossRef DOI. *Korea Information Processing Society*, 18(1), 1573-1575.
- Kim, Byung-Kyu, Kang, Mu-Yeong, Choi, Seon-Heui, Kim, Soon-young, You, Beom-Jong, & Shin, Jea-Do (2011). Citing behavior of Korean scientists on foreign journals in KSCD. *Journal of the Korean Society for Information Management*, 28(2), 117-133.
<http://dx.doi.org/10.3743/kosim.2011.28.2.117>
- Kim, Mi-Ra, & 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. <http://dx.doi.org/10.4275/kslis.2013.47.2.143>
- Seo, Tae-Sul, & Choi, Hee-Yoon (2011). Global dissemination of domestic scholarly journals using DOI and open access. *Journal of Information Management*, 42(4), 1-21.
<http://dx.doi.org/10.1633/jim.2011.42.4.001>