

국내 연구자의 학술 논문 CC 라이선스 채택 결정요인 분석*

An Analysis of Factors Influencing Creative Commons License Choices in Scholarly Articles by Korean Researchers

안병균 (Byoung-Goon An)**

이은지 (Eun Jee Lee)***

이혜진 (Hyejin Lee)****

초 록

학술 정보가 양질의 AI 학습데이터로 주목받기 시작하면서, 학술 정보의 CC 라이선스에 대한 관심이 높아지고 있다. 본 연구는 국내 연구자의 학술 논문 CC 라이선스 선택에 어떠한 요인이 영향을 미치는지 분석하는 연구이다. 이를 위하여 연구에서는 2020-2024년 출판된 한국 연구자의 하이브리드 학술지 오픈 액세스 논문 3,320편을 분석의 대상으로 선정하였다. 연구 지원 유형, 저자 구성, 교신저자 소속 기관 유형, 학문 분야, 저널 영향력에 따라서 CC 라이선스 선택에 나타나는 차이를 확인하고자 교차분석과 카이제곱 검정을 실시하고, 다항 로지스틱 회귀분석을 통해 영향 관계를 파악하였다. 그 결과, 외국인 저자가 포함될 경우 개방적인 CC 라이선스를 선택할 가능성이 커지며, 학문 분야에 따른 CC 라이선스 선택 차이도 드러났다. 또한 저널 영향력이 높은 학술지일수록 폐쇄적인 CC 라이선스를 선택하는 경향을 확인하였다. 본 연구는 국내의 CC 라이선스 제도 개선과 AI 학습데이터로서 학술 정보 활용 확대의 기초 자료로 활용될 것으로 기대된다.

ABSTRACT

As scholarly information increasingly gains attention as high-quality training data for artificial intelligence (AI), there is a growing interest in the Creative Commons (CC) licenses applied to academic publications. This study aims to analyze the determinants influencing the selection of CC licenses by Korean researchers for their academic papers. To this end, 3,320 open access articles published by Korean researchers in hybrid journals between 2020 and 2024 were selected for analysis. Cross-tabulation and chi-square tests were conducted to identify differences in CC license selection based on funding type, authorship composition, the corresponding author's institutional affiliation type, academic discipline, and journal impact. Furthermore, a multinomial logistic regression analysis was performed to determine the specific effects of these variables. The results revealed that the inclusion of international co-authors increased the likelihood of adopting open CC licenses. Conversely, articles published in journals with higher impact factors showed a tendency toward more restrictive CC licenses. The findings of this study are expected to serve as foundational data for improving the CC licensing policies in South Korea and expanding the utilization of scholarly information as AI training data.

키워드: CC 라이선스, CCL, 오픈 사이언스, 오픈 액세스, 학술 출판

CC License, Creative Commons License, open science, open access, scholarly publishing

* 이 논문은 2026년도 한국과학기술정보연구원(KISTI)의 기본사업으로 수행된 연구입니다

(과제번호: (KISTI) K26L2M1C1).

** 한국과학기술정보연구원 데이터큐레이션센터 선임연구원(rms0305@kisti.re.kr) (제1저자)

*** 한국과학기술정보연구원 데이터큐레이션센터 선임기술원(ejlee@kisti.re.kr) (공동저자)

**** 한국과학기술정보연구원 데이터큐레이션센터 책임연구원(hyejin@kisti.re.kr) (교신저자)

■ 논문접수일자: 2026년 2월 23일 ■ 최초심사일자: 2026년 3월 14일 ■ 게재확정일자: 2026년 3월 17일

■ 정보관리학회지, 43(1), 217-239, 2026. <http://dx.doi.org/10.3743/KOSIM.2026.43.1.217>

※ Copyright © 2026 Korean Society for Information Management

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits use, distribution and reproduction in any medium, provided that the article is properly cited, the use is non-commercial and no modifications or adaptations are made.

1. 연구 배경 및 목적

거대 언어 모델(Large Language Model, LLM)을 필두로 인공지능(Artificial Intelligence, AI) 기술이 발달하면서 전 세계적으로 인터넷 활용과 검색 양식이 변화가 생기는 한편, 생성형 AI의 오류와 할루시네이션이 문제가 되고 있다. 이러한 문제를 최소화하고 AI 모델의 성능을 향상시키는 핵심 요소로 AI 학습데이터의 중요성이 강조되고 있으며, 이는 비단 최근의 추세가 아닌 인공지능의 개발과 동시에 시작된 논의이다(Budach et al., 2022). 학술 연구 성과, 특히 새로운 지식 정보를 제공하고 동료 심사라는 과정을 통해 검증이 되었으며 논리적 완결성 또한 갖추고 있는 학술 논문은 고품질 AI 학습데이터로 주목받고 있다.

인공지능이 발전함에 따라 학술 연구 성과의 활용 방식이 기계학습으로까지 확장되면서, 연구 성과의 재이용 가능성이 중요해지고 있다. 연구 성과가 AI 학습데이터로 활용되는 과정에서 정제나 토큰화 등 저작물의 복제와 변형 과정이 수반되기 때문이다. 이러한 복제와 변형 과정은 저작권 침해의 소지가 있으며, 따라서 명시적인 이용 허락 없이는 AI 학습데이터로 활용이 어렵다. 오픈 액세스(Open Access) 운동을 통해 누구나 무료로 논문에 접근할 수 있는 환경이 조성되고 있으나, 이것이 기계에 의한 자유로운 학습과 가공을 허용한다는 의미는 아니다. 오픈 액세스 논문 중에서도 적합한 Creative Commons License(CC 라이선스)가 적용된 논문만이 단순한 접근 및 다운로드에서 더 나아가 자유롭게 재사용될 수 있다(Creative Commons, 2025).

CC 라이선스는 저작권법에 따라 창작자가 자신의 저작물을 대중에게 사용할 수 있도록 허가하는 표준화된 방법이다(Creative Commons, n.d.). CC 라이선스가 부여된 저작물은 잠재적 이용자에게 사전적으로 이용 허락을 한 것으로 간주되기 때문에(Margoni & Peters, 2016; Vollmer, 2022), CC 라이선스상 허가되어 있다면 AI 모델 개발을 위해 수만 건의 논문을 학습시킬 때마다 개별 저작권자와 별도의 권리 협상을 할 필요가 없다. 학술 연구 성과만을 위한 라이선스는 아니지만, CC 라이선스는 오픈 액세스 학술지와 논문에 기본적으로 사용되며 연구 성과의 이용과 배포, 변형의 범위를 규정하는 핵심적인 장치로 활용된다. CC 라이선스는 세부 종류마다 차이가 존재하며, 이에 따라 AI 학습데이터로서의 활용 가능성이 달라진다.

AI 학습데이터 활용 측면에서는 CC 라이선스를 명시하는 것뿐만 아니라 어떤 유형의 라이선스를 선택하는지 또한 중요한 사안이지만, CC 라이선스에 대한 실증 연구는 미비하다. 이에 따라 본 연구에서는 국내 연구자의 CC 라이선스 선택 경향을 파악하고, 이러한 선택에 영향을 미치는 요소를 확인하고자 한다. 이를 통해 CC 라이선스 선택에 대한 연구자의 인식을 파악하고 개방적인 라이선스의 확산을 위한 정책적 개입 지점을 도출하며, 향후 오픈 사이언스 정책과 학술 연구 성과의 AI 학습데이터 논의에 기초 자료를 제공하고자 한다. 이를 위해서 연구에서는 연구 지원, 저자 구성, 교신저자 소속 기관, 학문 분야, 저널 영향력(Impact Factor, IF)을 CC 라이선스 선택에 영향을 미치는 연구 외적 요인으로 선정하고, 계량정보학적 연구를 진행하였다. 본 연구에서 수립한 연구 질

문은 아래와 같다.

- [질문 1] 연구 지원 유형은 CC 라이선스 선택에 어떠한 영향을 미치는가
- [질문 2] 저자 구성은 CC 라이선스 선택에 어떠한 영향을 미치는가
- [질문 3] 교신저자 소속 기관 유형은 CC 라이선스 선택에 어떠한 영향을 미치는가
- [질문 4] 학문 분야는 CC 라이선스 선택에 어떠한 영향을 미치는가
- [질문 5] 저널의 영향력은 CC 라이선스 선택에 어떠한 영향을 미치는가

2. 이론적 배경 및 선행연구

2.1 학술 논문의 AI 학습데이터 활용

AI 학습데이터가 AI 기술 개발의 핵심으로 논의되고 있는 가운데(이재훈, 2020), 학술 논문은 고품질의 AI 학습데이터로 주목받고 있다. 실제로 대형 출판사인 Taylor & Francis와 Wiley에서 각각 자사 콘텐츠를 생성형 AI 모델의 학습에 사용할 수 있도록 계약을 체결한 바 있다(Gibney, 2024). AI는 학습된 수천만 건의 논문을 기반으로 개별 지식에서 지식의 연결고리를 도출하고, 복잡한 과학적 문제의 해결 속도를 높인다(Wang et al., 2022). 인공지능이 학술 논문을 학습함으로써 전 세계적으로 지식 생태계가 확장될 수 있는 것이다.

한편, 학술 논문을 AI 학습데이터로 활용하는 것에 대한 비판적인 시각도 존재한다. 학술 논문을 학습한 AI 모델이 산출 결과에서 원 저작자의 출처를 표시하지 않음으로써 지식 창출

에 대한 기여를 인정하지 않는다는 지적이 있으며(Nature Editorial, 2024), 이처럼 출처 표시가 불명확해질 경우 아이디어의 지적 계보를 추적할 수 없는 문제가 발생할 수 있다는 우려도 존재한다(Decker, 2025). 지적 재산권 측면에서 학술 논문의 AI 학습데이터 활용에 대한 우려는 타당한 것이다. 그러나 이러한 우려로 인해 학술 논문을 AI 학습데이터로 활용하는 것을 제한한다면 AI 성과와 학술적 영향력 측면에서 문제가 발생할 수 있다. AI 학습데이터가 편향될 경우, AI가 생성하는 정보의 정확도가 감소하며 할루시네이션이 발생할 가능성이 높아진다(Huang et al., 2025). 또한 생성형 AI가 새로운 검색 방식으로 대두되는 가운데, AI 학습데이터에 포함되지 않은 연구는 이용자에게 응답으로 표출될 가능성이 낮아지게 된다(Aggarwal et al., 2024). 이용자가 생성형 AI를 사용하여 검색 혹은 질의를 하는 상황에서 학습데이터 활용이 제한된 문헌과 그 지식은 결핍으로 도출될 가능성이 감소하며, 논문의 접근 및 이용과 지식의 확산 측면에서 불리하기 때문에 학술적 영향력이 낮아질 수 있다는 것이다.

학술 논문의 생성형 AI 학습은 돌이킬 수 없는 세계적인 추세이며, 이를 제한했을 때 발생할 수 있는 문제도 자명하다. 따라서 논문의 학습데이터 활용을 제한하는 방향보다는 연구 성과 활용의 공정성을 보장하면서 개방성을 유지하는 방법을 모색할 필요가 있다.

2.2 학술 논문과 Creative Commons License

논문 출판 시 출판사는 일반적으로 저자에게

출판 조건으로서 저작권을 양도하거나 이전하도록 요구한다. 출판사는 저작권 양도를 통해 일반적으로 복제, 개작, 배포, 공중 전시, 공중 공연에 대한 전반적인 독점적 권리를 부여받는다(Mondal, 2025). 이러한 독점적 권리를 토대로 출판사는 접근 제한이나 유료 구독 모델을 활용하여 금전적인 이익을 추구하고 있다. 학술 저자는 여타 창작물의 저작권자와는 다르게 배타적 권리를 행사하기보다는 개방적인 조건에서 자신의 창작물인 학술 논문이 공유, 활용, 인용되기를 지향함에도 불구하고(Vollmer, 2022), 출판사에 저작권을 양도하는 현 체계에서는 지향점과 상반되는 학술 논문의 접근 제한이 발생하는 것이다. 출판사에 저작권을 양도함에 따라 저자가 저작물의 공유 능력을 상실하게 되는 문제도 발생한다. 저작권자는 저작물에 CC 라이선스를 부여할 수 있는 유일한 존재이기 때문에 저작권이 양도되면 CC 라이선스를 통해 저작물을 공유하는 것은 불가능해지는 것이다(Vollmer, 2022).

오픈 액세스는 온라인에서 학술 논문을 무료로 이용하는 것에서 더 나아가, 해당 논문을 최대한 자유롭게 재사용할 수 있도록 하는 것까지 포함되어야 하며(Berlin Universities Publishing, n.d.), 이를 위한 도구로 CC 라이선스가 이용되고 있다. 특히 최근 인공지능 기술의 발전으로 학술 논문의 활용 범위가 TDM(Text and Data Mining)과 AI 학습데이터로까지 확장됨에 따라, CC 라이선스는 기계 학습 측면에서의 활용까지 규정하는 핵심 요소로 작동하고 있다. 저작자에 의해 CC 라이선스가 부여된 저작물은 저작자와 별도의 권리 협상을 할 필요 없이 이용자 누구나 라이선스 조건에 따라 저작물을

이용할 수 있다(Vollmer, 2022). CC 라이선스를 통해 저자가 저작권을 유지하고 적절한 출처 표기와 인용이 보장되면서 학술 연구 결과는 더욱 자유롭게 활용될 수 있는 유연성이 부여되는 것이다(Michael & Carroll, 2013; Mondal, 2025). 다만 오픈 액세스와 CC 라이선스가 자동으로 긍정적인 결과를 보장하는 것은 아니며, 정책, 기술, 윤리, 경제적 맥락 속에서 신중하게 적용하고 확산시킬 필요가 있다(Vollmer, 2022).

〈표 1〉은 Creative Commons에서 제공하는 라이선스 구분별로 의무 사항과 제한사항을 표기한 것이다. 저작권자가 자신의 저작물을 퍼블릭 도메인으로 배포하는 CC 0 외에, CC 라이선스는 SA(Share Alike) 조건, NC(NonCommercial) 조건, ND(No Derivatives) 조건을 조합하여 6가지 옵션을 제안한다. 이들 중 어떤 옵션을 선택하느냐에 따라서 AI 학습을 비롯한 저작물의 활용 가능 범위가 달라진다. NC 조건이 있는 라이선스의 경우 상업적 이익이나 금전적인 보상이 없는 방식으로만 사용을 허용하므로 오픈 액세스 정의와 양립할 수 없다는 견해가 존재한다(Margoni & Peters, 2016). 특히 AI 분야에서는 영리와 비영리 연구의 경계가 모호하고 상업적 모델 개발이 활발하므로, NC 조건은 해당 논문이 대규모 학습 데이터 세트에서 배제되는 결과를 초래할 수 있다. ND 조건은 2차적 저작물을 제작 및 배포하지 못하도록 제한하며, 이는 다른 언어로의 번역이나 주석 추가, 논문에 포함된 그림 및 이미지의 수정이 불가능함을 의미한다(Mondal, 2025). AI 학습데이터 활용을 위한 전처리 과정인 정제, 토큰화 등도 '수정'으로 해석될 소지가 있으므로 ND 라이선스는 AI 학습데이터 활용을 저해하는 요

〈표 1〉 CC0 및 CC 라이선스 옵션별 의무사항 및 제한사항

구분	원작자 명시 의무	동일 조건 라이선스 부여 의무	상업적 목적 제한	변형 제한 (파생, 각색)
개방적 ↑ ↓ 제한적	CC 0			
	CC BY	필수		
	CC BY-SA	필수	필수	
	CC BY-NC	필수	제한	
	CC BY-NC-SA	필수	제한	
	CC BY-ND	필수		제한
	CC BY-NC-ND	필수	제한	제한

인이 될 수 있다. 마지막으로 모든 형태의 재사용을 허용하나, 재생산된 콘텐츠를 게시하는 경우 동일한 라이선스 하에 게시해야 한다는 조건이 있는 SA 유형은 무료 콘텐츠의 재생산을 보장하는 것으로 보여진다. 그러나 더 개방적인 라이선스로 출판할 의향이 있거나 서로 호환되지 않는 여러 저작물을 동시에 사용하는 경우에는 문제가 발생할 수 있다(Berlin Universities Publishing, n.d.). 따라서 학술 논문의 자유로운 접근과 재사용 뿐만 아니라 AI 학습데이터로서 활용 가능성을 극대화하기 위해 가장 권장되는 유형은 논문의 자유로운 접근과 활용을 보장하는 CC BY라고 할 수 있다.

CC 라이선스를 통해 학술 논문의 자유로운 이용을 보장하고자 일부 연구 지원 기관에서는 CC 라이선스 사용을 의무화하고 있다. 영국, 유럽, 호주에서는 연구 지원을 기반으로 진행된 연구의 성과물에 대해 CC 라이선스 사용을 요구하고 있으며, 특히 RCUK와 World Bank에서는 CC BY 라이선스 선택을 의무화하였다 (Australian Research Council, 2021; European Research Council, 2012; Peters 2012; UK Research and Innovation, 2023). 논문 출판시 개별 출판사에서 특정 CC 라이선스를 권장

하는 경우도 있으나, 출판사의 경우 상업적 이익 추구를 위해 제한적인 라이선스를 선호할 가능성이 있다. 또한 출판사에서 연구자에게 라이선스 선택권과 정보를 명확하게 제공하지 않는 경우도 다수 존재하므로(Berlin Universities Publishing, n.d.), 오픈 액세스를 위한 CC 라이선스의 효과적인 활용을 위해서는 지속적인 교육과 훈련이 요구된다(Vollmer, 2022).

2.3 학술 논문의 Creative Commons License 분석 연구

학술 논문의 CC 라이선스에 대한 관심사는 높아지고 있으나, CC 라이선스 선택 경향이나 요인을 조사한 연구는 매우 드물다. Frass et al.(2013)은 전 세계 연구자를 대상으로 오픈 액세스 출판에 대한 설문을 진행하였으며, 그 일환으로 라이선스 유형별 선호도를 조사하였다. 그 결과, 연구자들이 CC BY-NC-ND에 대해서는 항상 선택한다는 대답이 71%로 가장 높은 반면 CC BY에 대해서는 선택하지 않는 다거나 특정 조건에서만 선택한다는 응답이 85%에 달했다. 또한 전체 응답 중 52%가 CC BY를 가장 선호하지 않는 라이선스 유형으로

선택하였다. 이를 통해 연구자들은 개방적인 라이선스보다 폐쇄적인 라이선스를 선호하는 경향이 있음을 확인할 수 있었다.

김규환, 정경희(2017)는 국내 학술지의 오픈 액세스 현황을 파악하고자 KCI 등재지 1,890종을 분석하였으며, 그 일환으로 CC 라이선스 현황과 특성을 조사하였다. 연구를 통해 무료 DB에서 접근 가능한 국내 KCI 등재지 중 CC 라이선스가 부여된 학술지는 26%임이 확인되었으며, 학문 분야 간 편차도 크게 나타났다. 연구에서는 CC 라이선스가 부여된 논문의 90% 이상이 CC BY-NC를 선택하고 있다고 밝히면서, 연구자나 학술지 편집인의 오픈 액세스 인식 증진을 위한 노력이 요구된다고 주장하였다. 후속 연구로 정경희(2023)는 국내 학술지의 CC 라이선스 적용 여부와 유형, 학문 분야와 발행 주체에 따른 차이를 조사하고자 한국연구재단 등재지 2,097종을 대상으로 분석을 진행하였다. 그 결과, 분석 대상의 18%만 CC 라이선스를 적용하고 있었으며, 학문 분야에 따른 편차가 크게 나타났다. CC 라이선스 유형 중 등재지에 가장 많이 적용된 유형은 CC BY-NC로 85%에 달했으며, 이러한 경향은 모든 학문 분야에서 유사하게 나타났다. 더불어 CC 라이선스를 적용한 등재지 중 절반 이상이 상용 DB에서 유료로 판매되고 있는 문제를 지적하였다.

Hong and Lee(2024)는 한국에서 발행되는 과학 및 공학 분야 오픈 액세스 저널의 저작권 정책을 분석하였다. 연구에서는 국제 학술 출판 플랫폼을 사용하지 않는 국내 과학 및 공학 분야 오픈 액세스 저널의 대부분이 저자에게 저작권 양도를 요구하고 있음을 확인하였다. 저작권 양도를 위한 계약서를 공개하는 저널은

다수 존재하였으나, 적절하고 유의미한 저작권 관련 조항이 포함된 경우는 매우 드물었다. CC 라이선스와 관련해서는 학술지 정책으로 CC BY-NC 라이선스 조건을 채택하는 것이 일반적이며, 저자가 라이선스 조건을 선택할 수 있는 사례는 존재하지 않았다.

3. 연구 방법

3.1 분석 대상 선정

본 연구에서는 한국인 연구자가 하이브리드 학술지에 오픈 액세스로 논문을 출판할 때 CC 라이선스 선택에 영향을 미치는 요인을 확인하고자 한다. '한국인 연구자'의 범위는 한국 기관에 소속된 저자이며, '한국인 연구자 논문'의 범위를 한국인 저자가 한 명 이상 포함된 논문으로 설정함으로써 외국과의 협력 연구 유형에 따른 CC 라이선스 선택 차이를 확인하고자 하였다. 모든 논문을 오픈 액세스로 출판하는 골드 학술지에는 특정 CC 라이선스 선택 규정이 있는 경우가 다수 존재하며, 대부분의 논문이 CC BY로 선택하고 있는 것이 확인되었다. 따라서 연구자의 CC 라이선스 선택이 비교적 자유롭고 다양한 하이브리드 학술지만을 대상으로 제한하였다.

2020년부터 2024년 사이 출판되어 Web of Science에 등재된 한국인 논문의 수는 400,905편이며, 20,856편의 논문이 하이브리드 학술지를 통해 오픈액세스로 출판되었다. 20,856편 중 논문 출판량이 가장 많은 상위 50개 학술지의 논문을 선별함으로써 전체 논문 중 30%를 분

석하고자 하였다. 마지막으로, 학술지의 CC 라이선스 정책이 연구자의 CC 라이선스 선택에 미치는 영향을 고려하여 상위 50개 학술지 중 학술지 정책을 통해 특정 CC 라이선스만을 선택하도록 규정한 15개 학술지의 논문을 제외하였다. 그 결과 3,323건의 논문이 분석의 대상으로 선정되었다.

3.2 변수 선정

본 연구는 연구 지원 유형, 저자 구성, 교신저자 소속 기관 유형, 학문 분야, 저널 영향력이 CC 라이선스 선택에 영향을 미치는지 확인하고자 한다. 종속변수인 CC 라이선스 유형은 Creative Commons에서 제공하는 유형을 활용하였으며, 독립변수는 연구 진행에 영향을 미칠 수 있는 정책적 배경, 연구 환경, 서지적 특성을 고려하여 선정하였다.

최근 연구 지원을 받은 연구 성과에 대해 CC BY와 같은 개방적인 라이선스를 적용할 것을 의무화하는 사례가 증가하고 있다(Ola, 2013; Petrak et al., 2021). 따라서 연구 지원 수혜 여부와 국내외 구분이 CC 라이선스 선택의 핵심 기제로 작용할 것으로 판단하여 연구지원유형을 변수로 설정하였다. 연구가 진행된 환경을 고려하고자 저자 구성과 교신저자 소속 기관 유형을 요인으로 채택하였다. 국제공동연구의 경우 참여 규모가 커질수록 강력한 개방형 라이선스 의무화 정책의 영향을 받을 가능성이 증가하며, 대규모 연구에는 글로벌 가시성과 재사용을 극대화하려는 동기가 강하게 작용된다. 또한 논문을 투고하고 CC 라이선스에 영향을 받는 교신저자는 소속된 기관의 정책을 반영할

수밖에 없으며, 소속 기관의 정책은 기관의 성격과 유형에 따라서 상이할 것이다. 이처럼 국제공동연구 양상과 저자 소속 기관의 유형은 CC 라이선스 선택에 영향을 미칠 수 있음에도 불구하고 관련 선행연구가 미비하다. 따라서 국제공동연구에 따른 저자 구성과 교신저자 소속기관 유형을 독립변수로 설정하고 본 연구를 통해 관계성을 밝히고자 하였다.

국내 선행연구를 통해 학문 분야에 따라 CC 라이선스 선택 양상이 상이하다는 것이 확인되었으므로(김규환, 정경희, 2017; 정경희, 2023), 구체적인 CC 라이선스 옵션 선택에 있어서도 영향을 미치는지 확인하고자 한다. 더불어, 학술적 권위가 높을수록 저자가 오픈 액세스 출판은 할지라도 활용에 있어서 강한 제약을 하는지, 반대로 더욱 적극적인 활용을 선택하는지를 확인하고자 저널의 영향력(Impact Factor, IF)을 연구 외적 요인으로 선택하였다.

3.3 데이터 수집 및 처리

데이터 수집 범위를 결정하고 대상 논문의 일부 정보(저널명, DOI, 출판연도, IF, 연구 분야)를 수집하고자 Clarivate의 InCites를 활용하였다. InCites를 통해 수집된 정보에서 IF가 누락되고 다른 정보원을 통해서도 확인이 불가능한 경우, 인접 연도의 값으로 대체하는 방식을 택하였다. 저자 구성과 교신저자 소속 기관, 연구 지원 정보를 확인하기 위하여 Web of Science의 논문 데이터를 활용하되, 필요에 따라 개별 논문의 페이지와 논문 본문을 확인하였다. CC 라이선스는 Crossref의 API를 이용하였으며, DOI를 통해 논문을 식별하고 'license' 필드의

값을 수집하였다. Crossref API를 통해 수집되지 않은 라이선스 정보는 개별 논문 페이지를 통해 확인하였다.

5가지 독립변수 중 IF를 제외한 4가지 독립변수는 각 기준에 따라 분류되었다. <표 2>는 본 연구의 독립변수 분류를 정리한 것이다. 연구 지원 유형은 국적과 지원 기관, 연구 지원 개수 등 굉장히 다양한 요인들이 있으나, 본 연구에서는 국내와 국외, 국내의 복합과 지원 없음으로만 구분하였다. 복잡한 연구 지원 정보의 구분을 단순화함으로써 국내 지원만 받은 경우와 국외 지원이 있는 경우의 차이를 확인하고, 연구 지원이 전혀 없는 경우 나타나는 양상을 파악하고자 하였다. 저자 구성은 교신저자 뿐 아니라 연구에 참여한 모든 저자를 고려하고자 국내와 국외 구분을 중심으로 전월 국내, 국내 주도, 국외 주도, 국내의 협력 네 가지로 구분하였다. 본 연구의 분석 대상 특성상 모든 연구자가 외국 기관 소속일 수는

없으므로 전월 국외는 분류에서 제외하였다. 교신저자 소속 기관의 경우 Research Organization Registry(ROR)에 명시된 Organization types를 참고하되, ROR에 정보가 없는 경우 기관의 웹사이트를 확인하여 소속 기관의 유형을 구분하였다. InCites에서 제공하는 논문별 세부 연구 분야는 그 구분이 지나치게 다양하므로, Eurostat이 개발한 과학분야 분류인 FOS 2007(Eurostat, 2013)을 적용하여 6개 학문 분야로 그룹화하였다.

3.4 통계분석

통계 분석에 앞서 한국인 저자의 CC 라이선스 선택 분포와 연도별 변화 양상을 조사하여 수집된 CC 라이선스 데이터의 전반적인 현황을 파악하였다. 이후 교차분석과 카이제곱 검정을 통해 연구 외적 요인과 CC 라이선스 유형 간의 상호 연관성을 파악하고 각 요인에 따라서

<표 2> 독립변수 분류 기준

데이터	분류
연구 지원 유형	<ul style="list-style-type: none"> - 국내: 한국에 있는 기관에서만 모든 연구 지원을 받은 경우 - 국외: 외국에 있는 기관에서만 모든 연구 지원을 받은 경우 - 국내외: 한국과 외국의 기관에서 모두 연구 지원을 받은 경우 - 지원 없음: 연구 지원을 받지 않았거나 연구 지원에 대한 언급이 없는 경우
저자 구성	<ul style="list-style-type: none"> - 전월 국내: 교신저자를 포함한 모든 저자의 소속 기관이 한국에 있는 기관인 경우 - 국내 주도(해외 참여): 교신저자의 소속 기관이 한국이며, 외국 기관 소속 공동저자가 있는 경우 - 해외 주도(한국 참여): 교신저자의 소속 기관이 국외이며, 한국 기관 소속 공동저자가 있는 경우 - 국내의 협력: 교신저자의 소속 기관에 한국과 외국이 모두 포함된 경우
교신저자 소속 기관 유형	<ul style="list-style-type: none"> - 교육 기관 - 기타(사기업, 정부 기관, 비영리 단체 등) - 연구 기관 - 의료 기관 - 복합
연구 분야	<ul style="list-style-type: none"> - 자연 과학 - 농업과학 - 기술 공학 - 사회과학 - 의학 - 인문학

CC 라이선스 선택 분포에 유의미한 차이가 있는지 검증하였다. 마지막으로 CC 라이선스 선택에 영향을 미치는 요인을 파악하기 위해 다항로지스틱 회귀분석을 진행하였다. 본 연구의 종속변수인 CC 라이선스 유형은 상호배타적이며 순서가 없는 명목형 범주이므로 선행회귀분석이나 이항 로지스틱 회귀분석이 불가능하며, 따라서 다수의 범주 중 하나가 선택되는 확률을 모델링하는 통계기법인 다항 로지스틱 회귀분석(유리하, 한경화, 2022)을 채택하였다. 종속변수는 CC 라이선스 중 가장 개방적인 옵션인 'CC BY'를 분석의 기준 범주로 설정하였으며, 독립변수의 기준 범주로 연구 지원 유형은 지원 없음, 저자 구성은 전원 국내, 교신저자 소속 기관은 교육 기관, 연구분야는 자연 과학으로 설정하였다. 회귀분석 모형의 적합도는 McFadden의 Pseudo R²를 통해 확인하였으며, 변수의 영향력은 오즈비(Odds Ratio)를 기준으로 해석하였다. 수집 데이터의 전처리와 기술분석은 Microsoft Excel에서 진행했으며, 카이제곱 검정과 다항 로지스틱 회귀분석은 SPSS 30을 활용하였다.

4. 연구 결과

4.1 독립변수 데이터 특성

〈표 3〉은 연구 분야, 연구 지원 유형, 저자 구성, 교신저자 소속 기관 유형 네 가지 독립변수의 각 분류별 빈도와 비율을 정리한 것이다. 연구 분야 기술통계 결과, 자연 과학 분야가 46.36%(1,539건)로 가장 많고 기술 공학(23.34%, 775건)과 의학(22.32%, 741건), 농업과학(7.98%, 265건) 순이었다. 사회과학 분야는 단 3건이었으며, 인문학 분야는 존재하지 않았다. 따라서 본 연구에서는 통계적으로 유의미한 표본 크기에 미치지 못한 사회과학 분야 3건의 논문을 분석 대상에서 제외함으로써 통계적 안정성을 확보하고자 하였다. 그 결과, 전체 분석 대상 논문의 수는 3,320건으로 조정되었다.

연구 지원 유형으로는 국내 지원(53.64%, 1,781건)이 가장 많았으며, 국외 연구 지원(16.05%, 533건), 국내의 연구 지원(15.18%, 504건), 연구 지원을 받지 않은 경우(15.12%, 502건)의 비율

〈표 3〉 독립변수별 세부 분류 빈도 및 비율

연구 분야		연구 지원 유형		저자 구성		교신저자 소속 기관	
자연 과학	1539 (46.36%)	국내	1781 (53.64%)	전원 국내	1839 (55.39%)	교육 기관	2150 (64.76%)
기술 공학	775 (23.34%)	국외	533 (16.05%)	국내 주도 (해외 참여)	436 (13.13%)	연구 기관	310 (9.34%)
의학	741 (22.32%)	국내외	504 (15.18%)	해외 주도 (국내 참여)	737 (22.20%)	의료 기관	43 (1.30%)
농업과학	265 (7.98%)	지원 없음	502 (15.12%)	국내외 협력	308 (9.28%)	그외	162 (4.88%)
사회과학	3					복합	655 (19.73%)
인문학	0						

은 큰 차이가 나타나지 않았다. 저자 구성은 연구 지원 유형과는 다소 다른 양상이 나타났다. 전원 한국인 저자로 구성된 경우가 55.39%(1,839건)로 가장 많은 것은 국내 연구 지원이 53.64%에 해당하는 것과 유사한 양상이지만, 국외 연구 지원은 15.18%에 그쳤던 반면 한국인 외 국적의 저자가 교신저자인 '해외 주도'는 두 번째로 많았다(22.20%, 737건). 국내외 협력은 9.28%(308건) 뿐이었다. 교신저자 소속 기관 유형에는 교육 기관이 64.76%(2,150건)로 가장 많고, 연구 기관이 뒤를 이었다(9.34%, 310건). 연구 분야에서는 의학이 22.32%를 차지한 반면 의료 기관이 차지하는 비중은 1.30%(43건) 뿐이었다.

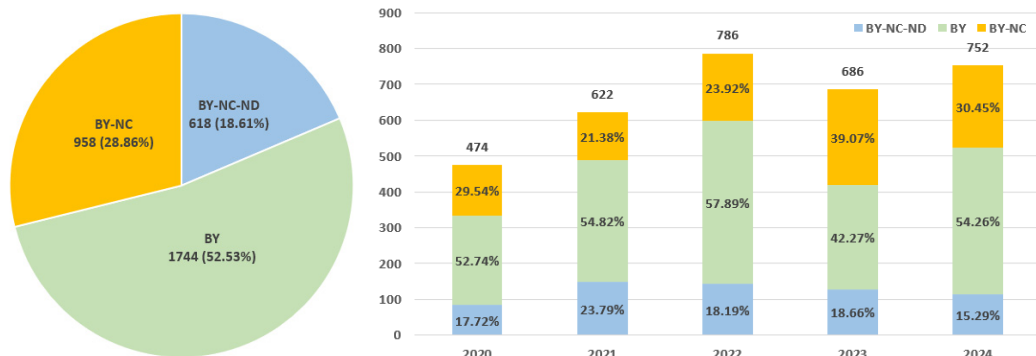
본 연구의 독립변수 중 하나인 IF에 대한 기술분석을 진행한 결과 평균은 7.67, 중앙값은 3.70으로 확인되었다. 이는 소수의 고인용지수 저널이 전체 평균을 상향시키고 있음을 의미한다. 또한 표준편차가 평균보다 큰 9.14이었으며, 최소값이 0.30, 최대값이 69.50으로 광범위한 분포를 드러냈다. 이를 통해 본 연구의 대상이 된 저널이 특정 수준에 국한되지 않고 다양한 영향력을 가진 저널을 포괄하고 있음을 파악할 수 있

었다. IF가 정규분포의 형태는 아닌 것으로 확인되었으나, 본 연구에서 진행되는 다항 로지스틱 회귀분석은 정규분포의 제약이 없으므로 IF를 독립변수에서 제외하거나 보정하지 않았다.

4.2 한국인 연구자의 CC 라이선스 선택 특성

〈그림 1〉은 한국인 저자의 CC 라이선스 옵션별 선택 비율과 시계열적 변화 양상을 분석한 것이다. 전체 3,320건 중 52.53%에 달하는 1,744건이 CC BY로 가장 높은 비율을 차지했으며, CC BY-NC(28.86%, 958건), BY-NC-ND(18.61%, 618건) 순이었다. CC BY-ND와 SA 조건이 포함된 옵션을 선택한 사례는 없었다. 이는 CC BY를 택하지 않는 연구자들의 고민이 영리 목적 이용을 제한할 것인지(BY-NC), 영리 목적 이용에 변경까지 불허할 것인지(BY-NC-ND)로 양분되고 있음을 보여준다.

다소간의 변동은 있으나 연도별 추이도 전체 비율과 유사한 양상을 보였다. 2023년을 제외한 모든 해의 CC BY가 항상 50% 이상이었으



〈그림 1〉 한국인 저자의 CC 라이선스 선택 특성

며, 2021년을 제외하면 CC BY-NC가 모든 연도에서 두 번째로 높은 비율을 차지했다. BY-NC-ND 유형도 20% 내외에서 등락을 반복하며 일정 수준을 유지하는 것을 확인할 수 있었다. 2021년의 경우 CC BY-NC-ND의 비율이 증가하면서 CC BY-NC보다 높아졌으나 2022년 다시 CC BY-NC-ND의 비율이 감소하면서 CC BY-NC가 두 번째로 높은 비율을 차지했다. 2023년에는 CC BY-NC 비율이 높아지면서 CC BY의 비율이 크게 감소하여 유일하게 50% 미만이 되었다. 2023년은 CC BY-NC와 CC BY의 비율에 큰 차이가 나타나지 않았으나, 2024년에는 다시 CC BY의 비율이 높아지면서 전체 비율과 유사한 양상을 보였다.

4.3 주요 변수별 CC 라이선스 선택 차이

연구 지원 유형, 저자 구성, 교신저자 기관 유형, 학문 분야 네 가지 독립변수와 종속변수인 CC 라이선스 선택 유형 간의 상호 관련성을 파악하고자 교차분석과 카이제곱 검정을 실시하였다. <표 4>는 연구 지원 유형에 따른 CC

라이선스 차이를 확인하기 위한 카이제곱 검정과 교차분석, 잔차분석 결과이다. 카이제곱 검정 결과, 연구 지원 유형과 CC 라이선스 선택 유형 간에는 통계적으로 매우 유의한 연관성이 존재하는 것으로 확인되었다($\chi^2 = 411.66, p < .001$). 즉, 연구 지원의 유무와 국외 지원 여부는 연구자의 라이선스 선택에 영향을 미치는 요인이라고 할 수 있다. 교차분석 결과, 국내 연구 지원만 있는 경우 CC BY 비율이 40.88%뿐인 반면, 국외 혹은 국내외 연구 지원을 받은 경우 각각 73.36%, 77.18%로 CC BY가 다른 CC 라이선스 옵션에 비해 월등하게 높은 비율을 나타내고 있었다. 잔차분석에서도 국내 지원을 받거나 지원이 없는 경우 CC BY보다 CC BY-NC를 선호하는 경향을 보이는 반면, 국외 혹은 국내외 지원을 받은 경우 반대로 CC BY를 선호하고 CC BY-NC를 비선호하는 경향이 드러났다.

저자 구성에 따라서 CC 라이선스 선택에 차이가 있는지 카이제곱 검정을 하고 교차분석을 진행했으며, 그 결과는 <표 5>와 같다. 카이제곱 검정 결과 $\chi^2 = 498.59$ 로 설명력이 매우 높

<표 4> 연구 지원 유형에 따른 CC 라이선스 차이 교차분석 및 카이제곱 검정

	BY	BY-NC	BY-NC-ND	합계	$\chi^2 = 411.66$ $p < .001$
국내	728 (40.88%) [-14.47]**	661 (37.11%) [11.30]**	392 (22.01%) [5.41]**	1781	
국외	391 (73.36%) [12.03]**	37 (6.94%) [-10.29]**	105 (19.70%) [-3.46]**	533	
국내외	389 (77.18%) [10.51]**	49 (9.72%) [-12.19]**	66 (13.10%) [0.70]	504	
지원 없음	236 (47.01%) [-2.69]**	211 (42.03%) [7.07]**	55 (10.96%) [-4.79]**	502	

1) () 비율
 2) [] 잔차분석 결과
 3) * $|z| > 1.96$ ($p < .05$), ** $|z| > 2.58$ ($p < .01$)

〈표 5〉 저자 구성에 따른 CC 라이선스 차이 교차분석 및 카이제곱 검정

	BY	BY-NC	BY-NC-ND	합계	$\chi^2 = 498.59$ $p < .001$
전원 국내	693 (37.68%) [-19.09]**	782 (42.52%) [19.37]**	364 (19.79%) [1.94]	1839	
국내 주도	256 (58.72%) [2.78]**	81 (18.58%) [-5.08]**	99 (22.71%) [2.36]*	436	
국외 주도	592 (80.33%) [17.13]**	49 (6.65%) [-15.08]**	96 (13.03%) [-4.42]**	737	
국내외 협력	203 (65.91%) [4.94]**	46 (14.94%) [-5.66]**	59 (19.16%) [0.26]	308	

- 1) () 비율
- 2) [] 잔차분석 결과
- 3) * $|z| > 1.96$ ($p < .05$), ** $|z| > 2.58$ ($p < .01$)

게 나타났으며, 통계적으로 유의미했다. 교차분석 결과, 국내 주도 유형, 국외 주도 유형, 국내외 협력 유형에서는 CC BY가 가장 높게 나타난 반면, 전원 국내 유형의 경우 CC BY-NC 유형을 선택한 비율이 42.52%로 가장 높았다. 해외 기관 소속 연구자가 교신저자인 국외 주도 유형에서 CC BY를 선택하는 경향이 가장 높게 나타났고(80.33%), 교신저자 소속 기관으로 국내와 국외 모두가 해당되는 ‘국내외 협력’ 유형의 CC BY 선택 비율이 65.91%, 교신저자는 국내 소속이지만 공동 저자에 해외 기관 소속 연구자가 포함된 경우 58.72%였다. 이는 해외 기관 소속 연구자의 참여도가 높아질수록 CC BY를 선택하는 비율이 높아진다는 것을 의미한다. CC BY-NC-ND 유형은 국내 주도 유형에서 가장 높았고(22.71%), 전원 국내 유형, 국내외 협력 유형, 국외 주도 유형 순으로 비율이 점차 낮아지고 있었다. 잔차분석에서도 동일한 경향이 나타난다. 특히 전원 국내 유형에서 CC BY를 기피하고 CC BY-NC를 강하게 선호하는 반면, 국외 주도 유형에서는 정반대로 CC BY에 대한 강한 선호가 나타

나고 CC BY-NC와 CC BY-NC-ND에 대한 기피가 나타났다. 이 또한 외국 기관 소속 연구자의 주도성이 높은 연구일수록 개방적인 CC 라이선스 유형을 선택한다는 것을 의미한다.

전반적인 저자 구성 외 교신저자의 소속 기관 유형이 CC 라이선스 선택에 차이를 발생시키는지 또한 확인하였다(〈표 6〉 참조). 카이제곱 검정 결과 $\chi^2 = 142.01$ 로 앞선 두 가지 설명력에 비하면 다소 낮으나 여전히 높은 설명력을 보였으며, 통계적으로도 유의미했다. 모든 기관에서 CC BY의 선택 비율이 가장 높게 나타났고, 연구 기관을 제외한 모든 기관이 50% 이상의 비율을 보였다. 그러나 교육 기관과 연구 기관의 경우 CC BY의 잔차가 음수로 나타나면서, 교육 기관과 연구 기관의 CC BY의 선호가 낮은 것으로 확인되었다. 교육 기관은 CC BY-NC를 선호하되 CC BY-NC-ND는 기피하는 현상을 보이면서 변형에 대한 제한이 상대적으로 적은 것으로 드러났다. 연구 기관과 의료 기관은 공통적으로 CC BY-NC-ND를 선호하는 것으로 나타났다.

〈표 7〉은 학문 분야에 따라 CC 라이선스 선

〈표 6〉 교신저자 소속 기관 유형에 따른 CC 라이선스 차이 교차분석 및 카이제곱 검정

	BY	BY-NC	BY-NC-ND	합계	$\chi^2 = 142.01$ $p < .001$
교육 기관	1101 (51.21%) [-2.07]*	725 (33.72%) [8.39]**	324 (15.07%) [-7.11]**	2150	
연구 기관	136 (43.87%) [-3.21]**	81 (26.13%) [-1.11]	93 (30.00%) [5.41]**	310	
의료 기관	24 (55.81%) [0.43]	5 (11.63%) [-2.51]*	14 (32.56%) [2.36]*	43	
기타	107 (66.05%) [3.53]**	40 (24.69%) [-1.20]	15 (9.26%) [-3.14]**	162	
복합	376 (57.40%) [2.79]**	107 (16.34%) [-7.89]**	172 (26.26%) [5.61]**	655	

- 1) () 비율
- 2) [] 잔차분석 결과
- 3) * $|z| > 1.96$ ($p < .05$), ** $|z| > 2.58$ ($p < .01$)

〈표 7〉 학문 분야에 따른 CC 라이선스 차이 교차분석 및 카이제곱 검정

	BY	BY-NC	BY-NC-ND	합계	$\chi^2 = 873.75$ $p < .001$
자연 과학	996 (64.72%) [13.07]**	182 (11.83%) [-20.13]**	361 (23.46%) [6.66]**	1539	
기술 공학	399 (51.48%) [-0.67]	259 (33.42%) [3.20]**	117 (15.10%) [-2.87]**	775	
의학	347 (46.83%) [-3.53]**	255 (34.41%) [3.79]**	139 (18.76%) [0.11]	741	
농업과학	2 (0.75%) [-17.60]**	262 (98.87%) [26.22]**	1 (0.38%) [-7.95]**	265	

- 1) () 비율
- 2) [] 잔차분석 결과
- 3) * $|z| > 1.96$ ($p < .05$), ** $|z| > 2.58$ ($p < .01$)

택에 차이가 나타나는지에 대한 카이제곱 검정 및 교차분석 결과이다. 카이제곱 검정 결과 통계적으로 유의미하며, $\chi^2 = 873.75$ 로 앞선 분석보다 더욱 높은 설명력을 보이면서 학문 분야에 따라 CC 라이선스에 차이가 발생함을 드러냈다. 자연 과학과 기술 공학의 경우 각각 64.72%, 51.48%로 CC BY가 50% 이상을 차지했으며, 자연 과학 분야의 경우 잔차분석에서도 CC BY에 대한 선호와 CC BY-NC 비선호가 강하게

나타났다. 자연 과학 분야는 CC BY 비율이 64.72%로 타 학문 분야의 CC BY와 비교해도 높은 반면, CC BY-NC(11.83%)보다 CC BY-NC-ND(23.46%)가 더 높게 나타났으며 잔차분석에서도 CC BY-NC-ND의 선호 경향이 드러났다. 즉, 자연 과학 분야는 가장 개방적인 형태를 선택하는 것이 일반적이나, 일부 연구에서는 가장 폐쇄적인 경향을 보인다는 것을 확인하였다. 의학 분야에서 가장 높은 비율을 차지

한 것은 CC BY(46.83%)이지만, CC BY에 대한 비선호가 나타났으며 오히려 CC BY-NC에 대한 선호가 나타났다. 교차분석 결과 농업과학 분야는 CC BY-NC가 98.87%로 극단적인 편중을 보였다. 이러한 편중은 다항 로지스틱 회귀분석에서 모형의 통계적 안정성을 저해할 수 있으므로, 회귀분석 단계에서는 농업과학 데이터를 제외하였다.

4.4 CC 라이선스 선택 영향 요인 분석

독립변수에 따라 CC BY 대비 CC BY-NC와 CC BY-NC-ND를 선택할 가능성이 어떻게 달라지는지를 확인하고자 다항 로지스틱 회귀분석을 진행하였다. 학문 분야 중 극단적인 편중 현상을 보인 농업과학 데이터를 제외한 3,055편

의 논문을 대상으로 회귀모형의 통계적 타당성을 검증한 결과, 통계적으로 유의하며 설명력도 높은 것으로 확인되었다(LLR $p < .001$, Pseudo $R^2 = 0.106$). 따라서 연구 지원, 저자 구성, 교신저자 소속 기관, 학문 분야, IF라는 독립변수가 연구자의 CC 라이선스 옵션 선택에 영향을 미치는지를 적절하게 설명하는 모형이라고 할 수 있다.

〈표 8〉은 독립변수에 따라 CC BY 대비 CC BY-NC를 선택할 가능성을 확인한 결과이다. 국내 연구 지원을 받는 경우 연구 지원을 받지 않을 때보다 CC BY-NC 대신 CC BY를 선택할 가능성이 더 낮아졌으며, 국외 혹은 국내의 연구지원의 영향은 통계적으로 유의미하지 않은 것으로 확인되었다. 저자 구성의 경우 전원 국내일 때에 비해 외국인 연구자가 교신저자 혹은

〈표 8〉 BY-NC 다항 로지스틱 회귀분석 결과

CC BY-NC		OR	B(Coef)	S.E.	p-value
Intercept		0.240	-1.426	0.184	0.000***
연구 지원 유형	국내	1.674	0.515	0.155	0.001**
	국외	0.776	-0.253	0.274	0.355
	국내외	0.981	-0.019	0.246	0.938
저자 구성	국내 주도	0.404	-0.907	0.167	0.000***
	국외 주도	0.264	-1.330	0.244	0.000***
	국내외 협력	0.582	-0.541	0.235	0.021*
교신저자 소속 기관 유형	연구 기관	1.118	0.111	0.170	0.513
	의료 기관	0.225	-1.492	0.507	0.003**
	기타	0.166	-1.794	0.409	0.000***
	복합	0.949	-0.053	0.143	0.713
학문 분야	기술 공학	2.569	0.943	0.129	0.000***
	의학	3.040	1.112	0.140	0.000***
저널 영향력(IF)		1.012	0.012	0.007	0.091

- 1) * $p < .05$, ** $p < .01$, *** $p < .001$
- 2) Reference Category: CC BY
- 3) Intercept: 연구 지원 없음, 전원 국내, 교육 기관, 자연 과학
- 4) Model Fit: Pseudo $R^2 = 0.197$, LLR $p < .001$

은 공동저자로 포함되어 있는 경우 CC BY-NC보다 CC BY를 선택하는 경향이 있는 것으로 나타났다. 특히 교신저자가 해외 기관 소속이고 국내 저자는 공동 저자로만 참여한 경우 CC BY-NC 대신 CC BY를 선택할 오즈가 3.788 (OR=1/0.264)배로 높아졌다. 교신저자 소속 기관에 따른 차이는 의료 기관과 기타 기관에서 확인되었다. 특히 의료 기관의 경우 CC BY 대비 CC BY-NC를 선택할 오즈비가 0.225로 나타나면서 CC BY-NC보다 CC-BY를 선택하는 경향이 드러났다.

학문 분야에 따른 차이도 확인되었다. CC BY 대비 CC BY-NC를 선택할 오즈비가 기술 공학은 2.569, 의학은 3.040으로 나타나면서 기술 공학, 의학 분야는 CC BY보다 CC BY-NC를 선택할 가능성이 높은 것으로 확인되었다. 비

교 기준이 된 자연 과학 분야는 다른 분야에 비해 CC BY-NC보다 CC BY를 선택할 가능성이 높은 것으로 나타났다. IF에 따라서 CC BY 대비 CC BY-NC를 선택할 오즈비는 통계적으로 유의미하지 않았다.

독립변수에 따라 CC BY 대비 CC BY-NC-ND를 선택할 가능성을 확인한 결과는 <표 9>와 같다. 연구 지원을 받지 않은 연구에 비해 국외에서만 지원을 받은 연구의 오즈비는 1.939로, CC BY 대비 CC BY-NC-ND를 선택할 가능성이 약 두 배로 나타났다. 저자 구성에서는 해외 기관 소속 연구자의 역할이 커질수록 CC BY-NC-ND보다는 CC BY를 선택할 가능성이 더 높게 나타났다. 저자로 참여한 연구자 전원이 국내 기관 소속일 때에 비해 공동 저자로 외국인 저자가 포함되어 있을 때 CC BY를 선

<표 9> BY-NC-ND 다항 로지스틱 회귀분석 결과

CC BY-NC-ND		OR	B(Coef)	S.E.	p-value
Intercept		0.318	-1.147	0.190	0.000***
연구 지원 유형	국내	1.380	0.322	0.175	0.066
	국외	1.939	0.662	0.244	0.007**
	국내외	0.781	-0.247	0.236	0.296
저자 구성	국내 주도	0.730	-0.315	0.153	0.040*
	국외 주도	0.245	-1.405	0.228	0.000***
	국내외 협력	0.455	-0.787	0.231	0.001**
교신저자 소속 기관 유형	연구 기관	1.921	0.653	0.158	0.000***
	의료 기관	1.548	0.437	0.359	0.224
	기타	0.575	-0.553	0.293	0.059
	복합	1.664	0.509	0.124	0.000***
학문 분야	기술 공학	0.726	-0.320	0.133	0.016*
	의학	1.014	0.014	0.141	0.921
저널 영향력(IF)		1.021	0.021	0.005	0.000***

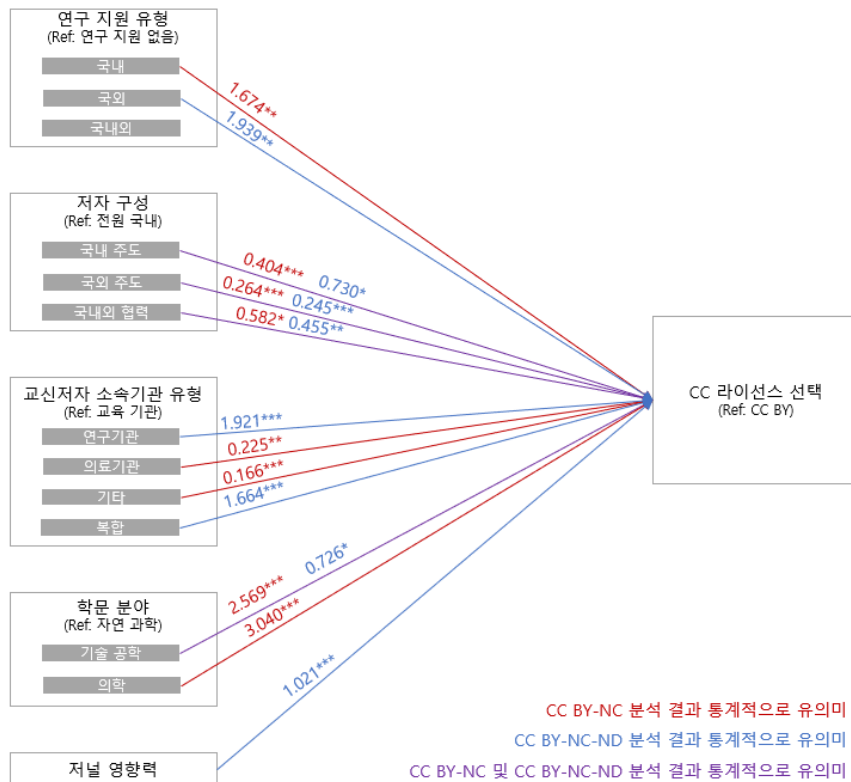
1) *p<.05, **p<.01, ***p<.001
 2) Reference Category: CC BY
 3) Intercept: 연구 지원 없음, 전원 국내, 교육 기관, 자연 과학
 4) Model Fit: Pseudo R²=0.197, LLR p<.001

택할 오즈비가 약 1.370(OR=1/0.730)이었으며, 교신저자가 해외 기관 소속 연구자로만 구성된 경우에는 4.082(OR=1/0.245)까지 증가하는 것으로 확인되었다.

교신저자 소속 기관 유형과 관련해서는 기관 유형에 따라 상이한 양상이 나타났다. 소속 기관이 연구 기관이나 기관 복합인 경우 CC BY-NC-ND를 선택할 오즈비가 각각 1.921, 1.664로 CC BY보다는 CC BY-NC-ND를 선택할 가능성이 높은 것으로 나타났다. 반면 의료 기관이나 기타 유형 기관의 경우 통계적으로 유의미하지 않았다. 학문 분야에 따라서는 기술 공학 분야만 통계적으로 유의미하게 나타났다.

자연 과학 분야에 비해 기술 공학 분야는 CC BY-NC-ND보다 CC BY를 선택할 오즈가 1.377(OR=1/0.726)배인 것으로 나타났다. 마지막으로 IF의 경우 1.021 오즈비를 드러내면서, IF가 1점씩 증가할 때마다 CC BY보다 CC BY-NC-ND를 선택할 가능성이 2.1% 증가하는 것으로 확인되었다.

본 연구를 통해 연구 지원 유형, 저자 구성, 교신저자 소속기관 유형, 학문분야, 저널 영향력 모두 CC 라이선스 선택에 영향을 미치는 것을 확인하였다. <그림 2>는 CC BY 대비 CC BY-NC와 CC BY-NC-ND에 대한 다항 로지스틱 회귀분석 결과를 도식화한 것이다.



<그림 2> 다항 로지스틱 회귀분석 결과

5. 논의

4장의 통계분석 결과를 토대로 독립변수별 분석을 진행하고 이를 기반으로 시사점을 도출하였다.

5.1 연구 지원과 개방

교차분석 및 카이제곱 검정 결과 연구 지원 유형에 따라 CC 라이선스 선택에 유의미한 차이가 확인되었다. 모든 연구 지원 유형에서 CC BY가 가장 높은 비중을 차지하고 있었으며, 국외와 국내의 지원 유형의 경우 70% 이상의 높은 비율을 보였다. 국외 지원과 국내의 지원을 받은 경우의 CC BY-NC 선택 비율은 10% 미만으로 국내 지원만 받은 경우보다 확연하게 낮았다. 이는 연구자들이 관행적으로 NC를 선택하는 경향이 남아있는 것과 더불어, 국외 연구 지원에 비해 국내 연구 지원은 CC BY에 대한 강제성이 약하기 때문에 나타나는 현상이다. 연구 지원이 없는 경우 유일하게 CC BY-NC에 대한 선호가 나타났다. 이를 통해 외부의 강요가 없다면 연구자의 자연스러운 선택은 상업적인 이용을 제한하는 것이라는 점을 확인할 수 있었다.

다항 로지스틱 회귀 분석 결과, 국내 연구 지원을 받은 경우 CC BY보다 CC BY-NC를, 국외 연구 지원을 받은 경우 CC BY-NC-ND를 선택할 가능성이 높은 것으로 나타났다. 펀딩을 받은 연구자들이 오히려 제한적인 라이선스를 선호하는 현상을 통해 현재 글로벌 오픈 사이언스 정책의 한계를 확인할 수 있다. 상당수의 연구비 지원 기관이 논문의 오픈 액세스 자

체는 의무화하고 있으나, 특정 개방형 라이선스의 채택까지는 강력하게 강제하지 않는 현행 정책의 한계에서 기인한 현상이다. 나아가 라이선스 강제 규정이 존재하더라도 이에 대한 지속적인 모니터링이나 위반 시의 제재가 이루어지지 않는 현실이 반영된 것으로 해석할 수 있다. 다만 본 연구의 대상은 하이브리드 저널의 오픈 액세스 논문이었기 때문에, 강력한 규정의 연구 지원을 받은 연구의 논문이 골드 혹은 다이아몬드 오픈 액세스 학술지에 우선적으로 투고되어 본 연구의 분석 대상에서 제외되었을 가능성을 배제할 수 없다.

5.2 국제 공동연구와 글로벌 라이선스 규범의 확산

저자 구성에 따른 CC 라이선스 선택 차이를 파악하기 위한 교차분석 및 카이제곱 분석 결과, 통계적으로 유의미한 관계가 있음이 확인되었다. 연구자 전원이 국내 기관 소속인 경우에만 CC BY-NC의 비율이 가장 높게 나타났으며, 외국 기관 소속 저자가 공동 저자로 포함된 경우, 공동 교신저자인 경우, 교신저자인 경우 순으로 CC BY의 비율이 높아지는 것으로 확인되었다.

다항 로지스틱 회귀분석 결과 전원 국내 기관 소속 연구자인 경우 CC BY-NC 혹은 CC BY-NC-ND를 선택할 가능성이 높게 나타났다. 국내 기관 소속 연구자가 교신저자이고 외국 기관 소속 공동저자가 존재하는 경우에는 저자 전원이 국내 기관 소속인 경우에 비해 CC BY-NC 혹은 CC BY-NC-ND를 선택할 가능성이 줄어드는 것을 확인할 수 있었다. 특히 국

외 기관 소속 연구자가 교신저자인 경우 전원 국내 기관 소속인 사례에 비해 CC BY를 선택할 가능성이 4.082배 높은 것으로 나타났다.

국외 기관 소속 연구자가 교신저자인 경우뿐만 아니라 공동연구자로 참여한 경우에도 CC BY를 선택할 확률이 1.370배 높아지는 것은 해외의 오픈 사이언스 문화가 공동연구를 통해 한국 연구 환경에 영향을 미치는 것이라고 할 수 있다. 국외 주도 연구와 국내외 협력 연구 모두 CC BY-NC-ND보다는 CC BY-NC를 선택할 가능성이 더 높은 것을 확인할 수 있었다. 국제 협력 연구의 경우 번역이나 데이터 공유 등 2차적 재사용이 빈번하게 일어나야 하므로, 이것을 제한하는 ND 라이선스의 선택을 지양하게 된 것이라고 할 수 있다.

저자 구성과 CC 라이선스 관계 분석 결과는 우리나라의 연구 성과 개방 문화가 아직은 국제적인 수준에 미치지 못했음을 드러낸다. 특히 전원 국내 기관 소속 연구자인 경우 폐쇄적인 라이선스를 선택할 가능성이 높아지는 것은 국내 연구 성과가 LLM의 학습데이터로 활용되거나 생성형 AI를 통해 확산되는 경로가 차단됨을 의미한다. 따라서 국내도 보다 개방적인 연구 문화를 정착시키고 이를 바탕으로 학술 논문의 개방을 활성화하기 위한 규범을 수립할 필요가 있을 것이다.

5.3 소속 기관의 특성에 따른 저작권 방어 기제

교신저자 소속 기관에 따른 CC 라이선스 차이의 카이제곱 검정 결과, 통계적으로 유의미한 것으로 나타났다. 교차분석에서 모든 기관

이 CC BY의 비율이 가장 높은 것으로 나타났으나, 연구 기관과 의료 기관의 경우 CC BY-NC-ND의 비율이 30%이상으로 높게 나타났으며, 잔차분석 결과에서도 CC BY-NC-ND를 선호하는 경향이 나타났다. 의료 기관의 경우 연구의 내용과 데이터가 개인정보, 민감정보, 의료 정보를 담고 있을 수 있으므로, 논문의 상업적 이용과 변형에 관한 제약이 크게 나타난 것이라고 할 수 있다.

다항 로지스틱 회귀분석 결과 의료 기관은 CC BY-NC보다는 CC BY를 선택할 가능성이 높게 나타났으며, 연구기관은 CC BY보다 CC BY-NC-ND를 선택할 가능성이 약 2배로 드러났다. 이와 같이 교신저자 소속 기관별로 CC 라이선스 선택 옵션이 상이하므로, 소속 기관의 특성을 고려하여 개방적인 라이선스 옵션 선택 확대를 위한 방안을 모색할 필요가 있을 것이다.

5.4 학문 분야에 따른 이용 제한

학문 분야에 따른 카이제곱 통계 결과 또한 높은 설명력을 지니며 통계적으로 유의미한 것으로 확인되었다. 자연 과학과 기술 공학, 의학은 CC BY의 비율이 가장 높았으며, 농업과학의 경우 CC BY-NC에 편향되어 있었다. 다항 로지스틱 회귀분석 결과, 의학 분야의 경우 CC BY보다는 CC BY-NC를 선택할 가능성이 약 3배 높은 것으로 나타났다. 변형 제한을 전혀 하지 않는 것은 아니지만, 변형 제한보다는 상업적 이용 제한에 좀 더 중점을 두는 경향이 있음을 확인할 수 있었다. 의학 분야에서는 CC BY보다 CC BY-NC를 선택할 가능성이 높게 나타난 반면, 의학 분야와 밀접한 관련이 있는

의료 기관 유형에서는 CC BY를 선택할 가능성이 더 높은 불일치 현상이 확인되었다. 이는 학문적 입장과 실용적 입장의 차이로 인해 발생한 현상으로 추정된다. 학문적 입장에서는 연구 성과의 무결성을 지향하며 이를 해칠 것으로 우려되는 변형 및 재가공을 금하지만, 의료 기관에서는 연구 성과를 임상 치료, 교육 자료, 가이드라인 등으로 활용해야 하므로 변형 및 재가공을 허용하는 방향으로 경향성에 차이가 드러난 것이다. 현재의 경향이 유지된다면 임상 현장의 논문은 가공과 AI 학습이 가능한 반면 그 기반이 되는 의학 논문은 가공 불가능으로 제한을 받게 되며, 이 경우 의료 전문 LLM이나 AI 개발에 제약이 생길 우려가 있다.

5.5 저널 영향력에 따른 보수성

다항 로지스틱 회귀분석을 통해 저널 영향력에 따라 CC 라이선스 옵션 선택에 차이가 발생하는지를 확인한 결과, IF가 1점 높아질 때마다 CC BY에 비해 CC BY-NC-ND를 선택할 가능성이 2.1% 증가하는 것으로 나타났다. 즉, 영향력이 높은 학술지에 출판될수록 연구자는 연구 성과를 개방하기보다는 활용을 제한하는 방식을 선택한다는 점이 드러났다. 이러한 결과는 권위 있는 연구일수록 학술 성과물의 변형 및 재가공 과정에서 발생할 수 있는 왜곡에 대한 연구자의 우려가 커진다는 것을 드러낸다. 또한 연구자가 저명한 학술지에 게재된 논문이 잠재적으로 상업가치가 있을 수 있다고 판단하는 경향이 있다는 것을 나타낸다.

고영향력 논문의 폐쇄성은 AI 학습 데이터

로서 고품질의 논문이 축적되기 어려운 구조를 만들게 된다. 이러한 구조가 고착화될 경우 미래의 AI는 최상위 연구 성과를 학습하지 못하고 상대적으로 접근이 용이한 저영향력 문헌 위주로 학습하게 될 것이다. 따라서 보다 개방적인 CC 라이선스 옵션을 선택할 수 있도록 장려하기 위해서는 연구자에게 내재되어 있는 변형 및 재가공에 대한 우려를 불식시키는 것이 필요하다. 또한 학술 논문의 개방이 AI 시대의 핵심 자원이 된다는 인식의 전환이 필요하며, 이를 확산시키기 위한 정책이 수립되어야 한다.

6. 결론

학술 정보가 양질의 AI 학습데이터로서 주목을 받고 있는 가운데, 학술 정보의 실질적인 활용이 주요한 이슈가 되고 있다. 오픈 액세스로 공개되어있는 학술 논문이라고 할지라도 그 활용 가능 수준은 라이선스에 따라 달라질 수 있으므로, 오픈 액세스 학술 논문의 라이선스를 규정하는 데 일반적으로 활용되는 CC 라이선스의 중요도가 높아지고 있다. 이에 본 연구는 국내 연구자가 하이브리드 학술지에 오픈 액세스로 논문을 게재하는 시점에서 선택하는 CC 라이선스에 어떠한 영향 요인이 있는지를 분석하고, 이를 기반으로 국내 학술 논문의 개방을 위한 시사점을 도출하고자 하였다.

이를 위하여 연구에서는 2020년부터 2024년까지 한국 기관 소속 연구자 1인 이상을 저자로 포함한 Web of Science 논문 3,320편을 분석하였다. 연구 지원 유형, 저자 구성, 교신저자 소속 기관 유형, 학문 분야, 저널 영향력이 CC

라이선스 선택에 영향을 미치는지를 확인하고자 카이제곱 검정과 교차분석을 수행하고 다항 로지스틱 회귀분석을 실시하였다.

연구 결과, 모든 변수가 통계적으로 유의미한 영향을 미치는 것으로 밝혀졌다. 펀딩 유형에 따라 라이선스 개방 정도가 상이했으며, 저자 구성에 있어서는 국외 연구자의 참여 여부와 정도에 따라 CC BY를 택할 가능성이 달라졌다. 교신저자 소속 기관 유형과 학문 분야에 따라 서로 다른 전략적 차이가 드러났다. 연구 기관에서는 CC BY-NC-ND를 선호하는 반면, 기술 공학 분야는 실질적 활용을 위해 ND 옵션은 선택하지 않는 경향을 보였다. 마지막으로, 저널 영향력이 높은 학술지에서 더 제한적인 옵션을 선택하는 것으로 나타났다.

저자 구성에 대한 분석을 통해 해외의 영향을 받는 경우, 상대적으로 개방적인 CC 라이선스 옵션을 선택하는 것을 확인하였다. 이는 한국과 협력하여 연구를 진행할 수준의 다른 국가에 비해 한국이 아직 CC 라이선스 규정을 비롯한 오픈 사이언스 제도를 정립하지 못했기 때문에 발생한 현상이다. 외부의 강제성이 부

재하면 개방보다는 제한을 선호하는 현재의 경향이 유지될 경우, 데이터 가공과 변환에 제약이 발생하고 AI 학습데이터 활용 및 AI 모델 개발에 제한이 생길 수 있으므로, 장기적으로 국내 AI 발전이 저해될 수 있다. 따라서 국내 연구 지원 기관에서 CC 라이선스 관련 정책을 수립함으로써 공유와 개방의 연구 문화를 조성하고, 기관과 학문 분야의 특성을 고려하여 라이선스 규정을 정립할 필요가 있을 것이다.

본 연구는 통계 모형의 안정성을 확보하기 위해 표본 수가 부족한 사회과학 분야와 특정 라이선스(CC BY-NC) 편중이 극심한 농업과 학 분야를 분석에서 제외하였다. 따라서 본 연구의 결과를 전 학문 분야로 일반화하기보다는, 과학기술 및 의학 분야의 출판 생태계에 국한하여 해석할 필요가 있다. 향후 인문 사회 분야까지 포괄하는 대상 범위에서 CC 라이선스 선택에 영향을 미치는 외부 요인을 분석하는 후속 연구를 하고자 한다. 일련의 연구는 국내 오픈 사이언스 제도의 수립과 AI 학습데이터로서 학술 정보 활용 확대의 기초 자료가 될 것으로 기대된다.

참 고 문 헌

- 김규환, 정경희 (2017). 국내 학술지의 오픈액세스 특성에 관한 연구: KCI 등재지를 대상으로. 정보관리학회지, 34(3), 251-267. <http://doi.org/10.3743/KOSIM.2017.34.3.251>
- 유리하, 한경화 (2022). 예측모형의 머신러닝 방법론과 통계학적 방법론의 비교: 영상의학 연구에서의 적용. 대한영상의학학회지, 83(6), 1219-1228.
- 이재훈 (2020). 인공지능 학습 데이터로서 공공연구데이터 활용 촉진을 위한 시론적 연구. 법과정제, 26(2), 95-128. <https://doi.org/10.36727/jlpr.26.2.202008.004>
- 정경희 (2023). KCI 등재지의 CCL 적용 현황 연구. 한국비블리아학회지, 34(2), 233-249.

- <http://doi.org/10.14699/kbiblia.2023.34.2.233>
- Aggarwal, P., Murahari, V., Rajpurohit, T., Kalyan, A., Narasimhan, K., & Deshpande, A. (2024). GEO: Generative engine optimization. In Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 5-16. <https://doi.org/10.1145/3637528.3671900>
- Australian Research Council (2021, September 1). ARC open access policy version 2021.1. Available: <https://www.arc.gov.au/sites/default/files/2022-06/Open%20Access%20Policy%20Version%202021.1.pdf>
- Berlin Universities Publishing (n.d.). Creative Commons Licences. Berlin Universities Publishing. Available: <https://www.berlin-universities-publishing.de/en/beratung/lizenzen/index.html>
- Budach, L., Feuerpfeil, M., Ihde, N., Nathansen, A., Noack, N. S., Patzlaff, H., Harmouch, H., & Naumann, F. (2022). The effects of data quality on machine learning performance. ArXiv preprint. <https://arxiv.org/abs/2207.14529>
- Creative Commons (2025, May 12). Frequently asked questions: Artificial intelligence and CC licenses. Creative Commons. Available: <https://creativecommons.org/faq/#artificial-intelligence-and-cc-licenses>
- Creative Commons (n.d.). About CC licenses. Creative Commons. Available: <https://creativecommons.org/share-your-work/ccllicenses/>
- Decker, S. (2025, April 15). Guest post – The open access - AI conundrum: Does free to read mean free to train? The Scholarly Kitchen. Available: <https://scholarlykitchen.sspnet.org/2025/04/15/guest-post-the-open-access-ai-conundrum-does-free-to-read-mean-free-to-train/>
- European Research Council (2012, June 1). Open access guidelines for researchers funded by the ERC. Available: http://erc.europa.eu/sites/default/files/document/file/open_access_policy_researchers_funded_ERC.pdf
- Eurostat (2013, June 4). Field of science and technology classification. Interoperable Europe Portal. Available: <https://interoperable-europe.ec.europa.eu/collection/eu-semantic-interoperability-catalogue/solution/field-science-and-technology-classification>
- Frass, W., Cross, J., & Gardner, V. (2013, March 1). Open access survey: Exploring the views of Taylor & Francis and Routledge authors. Taylor & Francis. Available: <https://files.taylorandfrancis.com/open-access-survey-march2013.pdf>
- Gibney, E. (2024, August 14). Has your paper been used to train an AI model? Almost certainly. Nature, 632, 715-716. Available: <https://doi.org/10.1038/d41586-024-02599-9>

- Hong, D. U. & Lee, J. Y. (2024). Copyright policies of science and engineering open access journals indexed in Science Citation Index Expanded or Scopus, published by Korean academic societies. *Science Editing*, 11(1), 62-72. <https://doi.org/10.6087/kcse.330>
- Huang, L., Yu, W., Ma, W., Zhong, W., Feng, Z., Wang, H., Chen, Q., Peng, W., Feng, X., Qin, B., & Liu, T. (2025). A survey on hallucination in large language models: Principles, taxonomy, challenges, and open questions. *ACM Transactions on Information Systems*, 43(2), 1-55. <https://doi.org/10.1145/3703155>
- Margoni, T. & Peters, D. (2016, March 10). Creative Commons licenses: Empowering open access. Social Science Research Network. Available: <http://doi.org/10.2139/ssrn.2746044>
- Michael, W. & Carroll, J. D. (2013). Creative commons and the openness of open access. *The New England Journal of Medicine*, 368(9), 789-791. <http://doi.org/10.1056/NEJMp1300040>
- Mondal H. (2025). Copyright and creative commons licenses in scholarly publishing: A practical guide for researchers. *Journal of Clinical Imaging Science*, 2025(15). http://doi.org/10.25259/JCIS_22_2025
- Nature Editorial (2024, August 27). AI firms must play fair when they use academic data in training. *Nature*, 632, 953. Available: <https://doi.org/10.1038/d41586-024-02757-z>
- Ola, O. (2013) Fundamentals of open access. *European Intellectual Property Review*, 36(2), 112-123. <http://doi.org/10.2139/ssrn.2382571>
- Peters, D. (2012, April 10). World Bank stakes leadership position by announcing open access policy and launching open knowledge repository under Creative Commons. Creative Commons. Available: <https://creativecommons.org/2012/04/10/world-bank-stakes-leadership-position-by-announcing-open-access-policy-and-launching-open-knowledge-repository-under-creative-commons/>
- Petrak, J., Skoric, L., & Macan, B. (2021). The impact of Plan S on scholarly journals from less developed European countries. *Croatian Medical Journal*, 62(1), 4-7. <http://doi.org/10.3325/cmj.2021.62.4>
- UK Research and Innovation (2023, May 24). UK Research and Innovation open access policy. UK Research and Innovation. Available: <https://www.ukri.org/publications/ukri-open-access-policy/uk-research-and-innovation-open-access-policy/>
- Vollmer, T. (2022). 16 Creative Commons and open access in an academic library: Implementation, tools, policy and education. In Coates, J., Owen, V. & Reilly, S. eds. *Navigating Copyright*

for Libraries: Purpose and Scope. Boston: De Gruyter Saur, 378-404.

<https://doi.org/10.1515/9783110732009-018>

Wang, H., Fu, T., Du, Y., Gao, W., Huang, K., Liu, Z., Chandak, P., Liu, S., Katwyk, P. V., Deac, A., Anandkumar, A., Bergen, K., Gomes, C. P., Ho, S., Kohli, P., Lasenby, J., Leskovec, J., Liu, T., Manrai, A., Marks, D., Ramsundar, B., Song, L., Sun, J., Tang, J., Veličković, P., Welling, M., Zhang, L., Coley, C. W., Bengio, Y., & Zitnik, M. (2023). Scientific discovery in the age of artificial intelligence. *Nature*, 620, 47-60.
<https://doi.org/10.1038/s41586-023-06221-2>

• 국문 참고문헌에 대한 영문 표기

(English translation of references written in Korean)

- Joung, Kyoung Hee (2023). A study on the current status of the adoption of Creative Commons licenses in Korean Citation Index journals. *Journal of the Korean Biblia Society for Library and Information Science*, 34(2), 233-249. <http://doi.org/10.14699/kbiblia.2023.34.2.233>
- Kim, Gyuhan & Joung, Kyoung Hee (2017). A study on the characteristics of open access in Korean scholarly journals: Focused on KCI journals. *Journal of the Korean Society for Information Management*, 34(3), 251-267. <https://doi.org/10.3743/KOSIM.2017.34.3.251>
- Lee, Jae Hoon (2020). A study on promoting the re-use of public research data as AI(Artificial Intelligence) learning data. *Law & Policy Review*, 26(2), 95-128.
<https://doi.org/10.36727/jlpr.26.2.202008.004>
- Ryu, Leeha & Han, Kyunghwa (2022). Machine learning vs. statistical model for prediction modelling: Application in medical imaging research. *Journal of the Korean Society of Radiology*, 83(6), 1219-1228.

