

생성형 인공지능 시대의 연구윤리 재정립: 주요 쟁점과 실천적 제언

김문정¹, 김도경^{2*}¹동아대학교 인문과학대학 철학생명의료윤리학과, 부교수²동아대학교 의과대학 의료인문학교실, 교수

Reframing Research Ethics in the Age of Generative Artificial Intelligence: Key Issues and Practical Proposals

Moon-jeong Kim¹, Do-kyong Kim^{2*}¹Associate Professor, Departments of Philosophy and Bioethics, College of Humanities, Dong-A University, Busan, Korea²Professor, Departments of Medical Humanities, College of Medicine, Dong-A University, Busan, KoreaReceived: Aug 10, 2025
Revised: Oct 23, 2025
Accepted: Oct 23, 2025***Corresponding author**Do-kyong Kim
Professor, Departments of Medical Humanities, College of Medicine, Dong-A University, Busan, Korea
Tel: +82-51-240-2743
E-mail: oliviak1211@gmail.comCopyright © 2025 Korean Society for Medical Ethics. This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.**Conflict of interests**

No potential conflict of interest relevant to this article was reported.

Acknowledgements

Not applicable.

Funding information

This work was supported by the NRF (National Research Foundation of Korea) Grant funded by the MOE (Ministry of Education) (NRF-2025J1A1A1A01022725).

Data availability

Upon reasonable request, the datasets of this study can be available from the corresponding author.

Abstract

This article explores how the traditional framework of research ethics must be redefined in response to the ethical challenges posed by the academic use of generative artificial intelligence (Gen AI). As Gen AI technologies become increasingly integrated across various stages of research—from study design and data analysis to manuscript writing—they offer notable gains in efficiency, while simultaneously expanding ethical gray areas. In particular, issues such as factual inaccuracies and hallucinations, violations of privacy and copyright, ambiguous authorship, and the deskilling of researchers are becoming increasingly entangled, posing serious threats to the credibility and integrity of academic work. To address these concerns, the article proposes four core ethical principles suited to the Gen AI era: truthfulness and explainability, respect for intellectual property, field-specific guidelines, and the cultivation of researcher competence. The central concern is not merely whether such technologies should be adopted, but rather how the identity and responsibility of researchers, and the sustainability of scholarly communities, can be ethically restructured in the face of transformative technological change.

Keywords: generative artificial intelligence; research ethics; integrity; accountability**I. 서론**

최근 생성형 인공지능(generative artificial intelligence, Gen AI)은 텍스트 생성, 이미지 및

ORCID 

Moon-jeong Kim
https://orcid.org/0000-0003-0291-4087
Do-kyong Kim
https://orcid.org/0000-0001-8639-1951

Author contributions

Conceptualization: Kim M.
Data curation: Kim M, Kim D.
Methodology: Kim M.
Investigation: Kim M.
Writing - original draft: Kim M, Kim D.
Writing - review & editing: Kim M, Kim D.

Ethics approval

Not applicable.

영상 제작, 소스 코드 작성 등 다양한 영역에서 비약적인 발전을 이루며, 디지털 전환을 가속화하는 핵심 기술로 부상하고 있다. 특히 OpenAI의 ChatGPT, Google의 Gemini, Meta의 LLaMA와 같은 거대 언어 모델(large language models, LLMs)은 인간 언어의 이해 및 생성 능력에서 전례 없는 성과를 보여주며, 인공지능 기술의 대표적인 진화 사례로 주목받고 있다. 이와 같은 기술의 고도화는 콘텐츠 생성 영역을 넘어 교육, 예술, 행정 등 사회 전반에 걸쳐 영향을 미치고 있으며, 학술 연구 분야 또한 그 예외가 아니다.

실제 연구 현장에서는 연구 설계, 데이터의 수집·처리·분석, 논문 작성, 참고문헌 요약 및 정리에 이르기까지, 연구 과정 전반에 걸쳐 Gen AI가 폭넓게 활용되고 있다[1]. 특히 반복적이고 시간 집약적인 업무의 부담을 줄임으로써 연구 효율성과 생산성을 높이는 유용한 도구로 인식되고 있다. 그러나 이러한 기술적 유용성에도 불구하고, Gen AI의 활용은 단지 실무적 편의성의 문제를 넘어, 연구윤리 전반에 중대한 도전 과제를 제기하고 있다.¹⁾

Gen AI가 산출하는 결과물은 그 출처와 정확성이 불확실할 수 있으며, 저작권 침해의 소지가 존재하고, 생성된 콘텐츠에 대한 책임 주체 역시 불분명한 경우가 많다. 아울러, 연구자가 직접 수행하던 분석과 창작 과정이 자동화되면서 연구 역량의 약화, 이른바 탈숙련화(deskilling) 문제가 대두되고 있다. 특히, Gen AI가 생성한 텍스트나 아이디어가 창작물로 인정될 수 있는지, 혹은 표절의 범주에 속하는지에 대한 판단은 기존 연구윤리 규범의 경계를 모호하게 만들고 있으며, 연구 결과에 대한 책임(accountability) 문제 역시 구조적 재검토를 요구한다.

본 논문은 Gen AI의 학술적 활용이 불러일으키는 윤리적 쟁점을 분석하고, 이에 대응하는 연구윤리의 새로운 기준을 모색하는 데 목적이 있다. 이를 위해 먼저 Gen AI 관련 주요 윤리 문제—예컨대, 환각(hallucination) 현상, 개인정보 및 저작권 침해, 저작성(authorship) 논란, 탈숙련화 등을 중심으로 논의하고자 한다. 이어서 Gen AI를 연구에 통합적으로 활용하기 위해 연구자에게 요구되는 윤리적 태도와 책임 기준을 도출하며, 마지막으로 연구자 개인뿐만 아니라 학문 공동체 및 제도적 차원에서의 실천적 대응 방안을 제시함으로써, Gen AI 시대에 걸맞은 연구윤리 체계의 재정립 방향을 탐색하고자 한다.

II. Gen AI(Artificial Intelligence)의 연구 활용에 따른 주요 윤리 쟁점

본 장에서는 Gen AI의 연구 활용 과정에서 제기되는 핵심적인 윤리 문제들을 고찰하고자 한다. Gen AI는 기술 발전과 더불어 학술 환경 전반에 빠르게 확산하며 새로운 연구 도구로 평가되고 있으나, 그 활용은 동시에 다면적인 윤리 쟁점을 수반한다. 구체적으로는 생성 결과의 신뢰성 결여, 개인정보 침해 가능성, 저작권 및 지식재산권의 위반, 연구 책임의 불분명성 등이 대표적이다. 이에 Gen AI의 연구 활용과 관련하여 주목해야 할 주요 윤리 문제들을 네 가지 범주로 구분한 뒤, 각각의 논점을 분석함으로써 윤리적 대응 전략의 기초를 제공하고자 한다.²⁾

1) Panda & Kaur[1]는 Gen AI가 학문 연구와 교육의 전 과정에 걸쳐 미치는 영향을 종합적으로 고찰하면서, 그 활용 영역을 문헌 탐색, 데이터 시각화, 콘텐츠 생성, 표절 탐지, 데이터 분석, 언어 향상(교정), 투고 저널 추천 등 일곱 가지 범주로 체계화하였다. 이들은 Gen AI가 연구 전 단계에서 생산성과 효율성을 높이고 학문적 접근성을 확장하는 기회를 제공한다고 평가하면서도, 동시에 사실과 다른 응답(hallucination)의 생성, 비판적 사고의 약화, 표절 및 저자 자격 논란, 기존 평가 체계의 붕괴, 제도적·윤리적 규범의 미비와 같은 중대한 도전과제를 지적한다.



1. 환각과 정보 오류

Gen AI는 대규모 데이터를 기반으로 통계적 규칙성과 언어적 패턴을 학습함으로써, 외견상 그럴듯한 텍스트를 산출하는 기능을 수행한다. 그러나 이러한 생성 과정은 정보의 진위나 논리적 타당성을 판별하는 인식적 작용에 기반하지 않으며, 확률적 연관성에 따른 기계적 계산의 결과에 불과하다. 이로 인해, 실제로 존재하지 않는 사실을 포함하거나 비논리적이고 황당한 내용을 생성하는 ‘환각’ 현상이 발생하게 되며, 이는 특히 학술 연구의 신뢰성과 진실성을 심각하게 훼손할 수 있는 요인으로 작용한다.³⁾ 즉 Gen AI의 산출물을 비판 없이 학술적 글쓰기나 발표에 반영할 경우, 왜곡된 정보가 학문적 담론 속에 무비판적으로 수용될 수 있으며, 이는 학술 생태계 전반에 부정적 영향을 미칠 수 있다.

환각 현상은 단순한 시스템 오류가 아닌, 학습 데이터의 성격 및 처리 방식에 기인하는 구조적 한계에서 비롯된다[2]. Gen AI는 통제되지 않은 온라인 자료를 포함한 방대한 데이터를 학습하는데, 이 과정에는 검증된 정보뿐 아니라 사실 여부가 불명확하거나 논쟁의 여지가 있는 내용, 심지어 명백히 허위이거나 편향된 정보도 포함되어 있다[3]. 정보량의 증가로 오류가 상쇄될 가능성이 일부 존재하나, 본질적으로 검증되지 않은 데이터가 학습에 포함될 수밖에 없는 한계로 인해 환각 발생은 피할 수 없는 문제로 남는다. 더욱이 Gen AI의 대표적 구현 형태인 LLMs은 사용자의 입력에 대응하여 통계적으로 가장 자연스러운 응답을 생성하는 데 초점을 맞추는데, 그 과정에서 사실관계를 판별하거나 검증하는 기능은 본질적으로 결여되어 있다.⁴⁾ 이러한 생성 방식은 외형상 완결성 있는 결과물을 제공하면서도, 내용적으로는 오류와 허위 정보를 포함할 가능성을 항상 내포한다.

실제로 이와 같은 한계는 여러 사례를 통해 문제화된 바 있다. 예컨대 미국의 한 변호인은 Gen AI가 생성한 허위 판례를 인용하여 법원에 제출하였다가 징계를 받았으며[4,5], 의료 분야에서는 존재하지 않는 논문을 기반으로 진단 또는 치료의 근거를 삼은 사례가 보고된 바 있다[6]. 이러한 사례는 Gen AI의 환각이 현실적이고 구체적인 피해를 초래할 수 있음을 보여

2) 필자는 Gen AI의 학술 연구 활용에서 제기되는 윤리적 쟁점을 ‘연구 전주기(research lifecycle)’의 단계와 연계하여 네 가지 범주로 구분하였다. 첫째, 연구 기획 및 설계 단계에서는 연구자의 역할 분담과 책임 귀속의 문제가 대두되며, 이에 따라 저자 자격 논란과 책임이 핵심 윤리적 쟁점으로 제기된다. 둘째, 데이터 확보 단계에서는 데이터의 출처와 적절성, 수집 과정의 투명성이 중대한 과제로, 프라이버시 침해와 저작권 침해 문제가 두드러진다. 또한 이 단계에서 불완전하거나 편향된 데이터가 확보될 경우 이후 연구 전 과정에 영향을 미치므로, 환각과 정보 오류의 잠재적 원인 역시 이 단계에서 발생한다. 셋째, 데이터 활용 단계에서는 인공지능이 확보된 데이터를 처리·분석하는 과정에서 결과의 신뢰성과 정확성이 검증되며, 이때 환각과 정보 오류가 구체적으로 드러난다. 넷째, 결과 도출 및 논문 작성 단계에서는 인공지능 의존이 연구자의 비판적 사고와 창의성을 약화시킬 위험이 있으며, 이는 연구자 탈숙련과 연구 역량 저하라는 쟁점으로 연결된다. 따라서 이 네 가지 분류는 연구 전주기의 핵심 단계에 각각 대응하는 윤리적 기준에 의해 체계화된 것이라 할 수 있다.

3) Gen AI는 전통적인 인공지능이 단순히 기존 데이터를 분석하거나 예측하는 수준을 넘어서, 새로운 합성 데이터를 창출하고, 문장 및 내용을 생성·보정하며, 데이터 증강(data augmentation)까지 수행할 수 있는 능력을 지닌다. 그러나 이러한 적극적 기능은 사실과 다른 정보, 즉 환각을 생성할 위험을 본질적으로 내포하며, 이는 연구윤리적 차원에서 위조와 변조 문제와 긴밀히 연결된다. 합성 데이터는 실제로 존재하지 않는 자료를 만들어내어 연구 결과를 왜곡할 수 있고, 데이터의 보정이나 가공 과정은 연구자가 의도하지 않은 변형을 발생시킬 수 있기 때문이다. 따라서 생성형 인공지능의 환각 현상은 단순한 기술적 오류로 축소할 수 없는 문제이며, 연구 활동의 진실성과 신뢰성을 위협하는 윤리적 위험으로 파악되어야 한다.

4) 구글 AI 윤리 연구팀을 이끌었던 팀니트 게브루(Timnit Gebru)는 공동 저자들과 함께 『On the dangers of stochastic parrots: can language models be too big?』라는 논문을 통해, ChatGPT와 같은 LLMs의 훈련 및 학습 방식에 내재된 한계를 비판적으로 분석하였다. 이들은 LLMs이 언어를 그럴듯하게 생성할 수는 있지만, 실제로는 언어를 이해하지 못한 채 통계적 확률에 기반한 모방을 수행하는 ‘확률적 앵무새(stochastic parrot)’에 불과하다는 점을 지적하였다[3].

준다.⁵⁾

이와 같은 문제를 완화하기 위한 기술적 대응도 다양하게 제안되고 있다. 신뢰 가능한 데이터베이스를 기반으로 한 생성 응답의 교차 검증, AI의 산출물을 다른 언어모델이나 전문가가 재검토하는 다중 확인 체계, 인간 피드백을 활용한 강화 학습(reinforcement learning from human feedback, RLHF), 생성 내용에 대한 출처 강제 표기 등이 그 예이다[7]. 그러나 이러한 조치들은 환각 현상을 일정 부분 줄일 수는 있어도, Gen AI의 작동 원리 자체가 통계적 추론에 기반하고 있는 한, 오류 가능성을 완전히 제거할 수는 없다. 이는 Gen AI가 내포하는 구조적 불확실성으로, 결과물의 진실성 확보에 본질적인 한계를 가질 수밖에 없다는 사실을 의미한다.

따라서 연구자는 Gen AI의 결과물을 신뢰 가능한 자료로 간주해서는 안 되며, 생성된 내용에 대해 사실 여부, 출처의 적절성, 논거의 정당성 등을 엄밀히 검토할 책임이 있다. 특히 논문, 연구보고서, 정책 제안서 등과 같은 학술적·공적 문서 작성 시에는 AI의 응답이 오류를 포함하고 있을 가능성을 전제로, 철저한 사실 확인과 검증 절차가 반드시 요구된다. 이는 단지 연구자의 윤리적 책무에 그치지 않으며, 학문적 정직성과 연구진실성(research integrity)을 유지하기 위한 불가결한 전제라고 할 수 있다.

2. 저작권 및 프라이버시 침해

Gen AI는 웹상에 공개된 다양한 유형의 콘텐츠—예컨대 뉴스 기사, 블로그 게시물, 위키 문서, 전자 학술지, 이미지, 영상 등—를 학습 자료로 삼아 언어적 및 시각적 생성 능력을 고도화하고 있다. 그러나 이와 같은 대규모 학습 과정에는 개인정보 및 저작권 보호 대상 자료가 무분별하게 포함될 가능성이 내재되어 있으며, 이는 연구자에게 프라이버시 보호와 저작권 준수라는 이중의 윤리적 책임을 부과한다.⁶⁾

먼저, 개인정보와 관련된 윤리적 문제는 데이터 수집 방식에서 비롯된다. 웹 기반 정보의 상당수는 정보 제공자의 명시적 동의 없이 수집되며, 설령 형식적 동의가 이루어졌다고 하더라도, AI 학습이라는 특수한 활용 목적에 대한 충분한 설명이나 이해를 전제로 한 것이 아닌 경우가 많다[8]. 특히 대다수 사용자들이 광범위한 약관을 면밀히 검토하지 않은 채 단순히 ‘동의’ 버튼을 클릭하는 관행은, 이를 ‘충분한 정보에 근거한 동의(informed consent)’로 간주할 수 있는지에 대한 법적·윤리적 논쟁을 불러일으킬 수 있다[9]. 이로 인해 Gen AI가 학습한 자료 중에는 윤리적 정당성이 결여된 개인정보가 포함될 가능성이 있으며, 연구자가 이러한 데이터를 기반으로 생성된 결과물을 비판 없이 활용할 경우, 의도하지 않더라도 프라이버시 침해에 가담하게 될 위험이 존재한다.⁷⁾

저작권 문제는 더욱 복잡하고 다층적인 양상을 보인다. Gen AI는 기존의 저작물을 학습한

5) 지난 5월 정보통신기획평가원이 발표한 주요 AI 모델의 분야별 환각 감지 정확도 분석에 따르면, AI는 전문성이 요구되는 분야일수록 비논리적이거나 사실과 다른 응답을 생성할 가능성이 더 높은 것으로 나타났다[4].

6) Gen AI는 방대한 데이터셋을 기반으로 학습하는 과정에서 저작권이 보호된 자료를 무단으로 활용할 가능성이 크며, 그 결과물 역시 기존 저작물의 표현이나 구조를 변형·전용하여 저작권 침해의 위험을 내포한다. 이러한 특성은 학문적 맥락에서 연구자가 AI 생성물을 출처 표기 없이 활용할 경우 곧바로 연구윤리의 핵심 위반인 표절로 이어질 수 있다. 즉, 저작권 침해가 법적 차원의 문제라면, 표절은 학문 공동체의 신뢰를 근본적으로 훼손하는 윤리적 문제로, 두 영역은 생성형 인공지능 활용에서 긴밀히 연결된다. 따라서 연구자는 AI 산출물의 출처와 생성 과정을 투명하게 드러내고, 적절한 검증 및 인용 원칙을 준수해야 하며, 더 나아가 학문 공동체와 제도적 차원에서 저자 자격, 책임성, 인용 규칙 등을 새롭게 재정립하는 노력이 요구된다.

후, 이를 조합하거나 재구성하여 새로운 콘텐츠를 생성하는데, 이 과정에서 생성된 결과물은 원저작물과 구조적·표현적으로 유사한 형식을 가질 수 있다[10]. 특히 논문 요약, 표 및 그래프 작성, 시각 자료 구성 등에서 AI가 학습한 기존 텍스트나 이미지를 의식하지 못한 채 재현할 경우, 이는 저작권 침해 또는 표절(plagiarism)로 간주될 수 있다. 더욱이, 생성된 결과물을 이차적 저작물로 활용하는 과정에서 원저작물에 대한 적절한 출처 표기나 권리 확인이 이루어지지 않을 경우, 연구의 신뢰성을 심각하게 훼손할 뿐만 아니라, 저작권 침해로 인한 법적 책임이 발생할 수도 있다.

이러한 문제는 Gen AI의 작동 방식 자체가 출처 추적에 구조적 한계를 지니고 있다는 점에서 더욱 복잡해진다. 생성 과정에서 활용된 데이터의 경로가 불투명하거나 비가시적인 경우가 많기 때문에, 연구자가 결과물의 정확한 출처를 식별하는 데 현실적인 제약이 따를 수밖에 없다. 그럼에도 불구하고, 학술 연구의 윤리적 관점에서 생성된 정보라 하더라도 인용 규범과 출처 명시 원칙은 여전히 엄격하게 요구되며, 이를 회피하거나 간과하는 것은 연구의 정당성과 신뢰성을 저해할 수 있다. 이러한 조치는 단순한 기술적 투명성을 넘어서, 연구자의 진정성과 학문 공동체로부터의 신뢰 확보를 위한 필수적 조건이다.

결론적으로, Gen AI 시대의 프라이버시와 저작권 문제는 단순한 도구적 효율성의 문제가 아니라, 연구자의 윤리적 통찰과 자율성, 그리고 투명한 연구 수행에 기반한 실천의 문제로 귀결된다. 연구자는 인공지능이 제공하는 편의성에 무비판적으로 의존하기보다는, 정보 취득과 활용의 정당성 여부를 법적·윤리적 기준에 따라 자율적으로 점검할 수 있어야 하며, 학술적 산출물의 모든 구성 요소에 대한 책임 있는 태도를 견지해야 한다.

3. 저자 자격 논란과 책임의 문제

Gen AI의 활용이 학술 연구 및 출판 전반에 걸쳐 확산됨에 따라, 그 산출물에 대한 저자 자격 인정 가능성이 새로운 윤리적 쟁점으로 대두되고 있다. 이 문제는 단순한 명의 표기의 여부를 넘어, 학술 저작물에 요구되는 책임성과 해명 가능성이라는 근본적인 윤리 기준과 밀접하게 연결되어 있다.⁸⁾

현재 국제 학술 출판계는 Gen AI를 논문의 정식 저자로 인정하지 않는 일관된 입장을 표명하고 있다[11-16]. 국제학술지출판윤리위원회(COPE), 국제의학학술지편집인위원회(ICMJE), 세계의학편집인협의회(WAME) 등 주요 기관들은 공통적으로 “저자는 연구 결과에

7) ‘충분한 정보에 근거한 동의’는 본래 의료 및 임상 연구의 맥락에서, 환자 또는 피험자가 자신의 건강과 신체에 중대한 영향을 미치는 의사결정에 능동적으로 참여할 수 있도록 보장하기 위해 발전된 윤리적·법적 개념이다[9]. 그러나 정보기술의 비약적인 발전과 함께 온라인 서비스, 디지털 플랫폼, 인공지능 활용 환경 등에서도 개인의 정보결정권과 자기결정권이 핵심적인 쟁점으로 부상하면서, 이 개념은 점차 비의료 영역으로까지 확대 적용되고 있다. 이 글에서는 임상 현장에서 요구되는 엄격한 기준과는 차이가 있음을 전제하면서도, 동의 과정의 실질성과 자율성 보장의 중요성을 공유한다는 점에서, ‘충분한 정보에 근거한 동의’ 개념을 웹 기반 환경에서도 유효한 규범적 분석틀로 일반화하여 적용하는 것이 타당하다고 본다.

8) ‘Gen AI의 학술적 활용은 기존 연구윤리 담론에서 저자 자격과 책임의 문제를 새롭게 부각시킨다. 특히 인공지능이 작성 과정에 참여한 연구 산출물을 저자성의 범주에 포함시킬 수 있는지에 대한 논의는 전통적으로 연구부정행위로 규정되어 온 부당한 저자표시 문제와 밀접히 연결된다. 국제적 기준에 따르면 저자는 연구의 설계, 자료 해석, 논문 작성과 같은 실질적 기여를 수행하고, 연구 전 과정에 대한 책임성을 분명히 가져야 한다. 이에 따라 Gen AI는 독립적 의사결정 능력이나 도덕적 책임성을 지니지 못하므로 저자로 등재될 수 없으며, 단지 연구를 보조하는 기술적 수단으로만 활용이 허용된다. 그럼에도 불구하고 이를 저자로 표기하는 행위는 연구의 책임 구조를 왜곡하고 연구자의 설명 책임을 회피하는 결과로 이어지므로, 연구 부정행위에 해당된다고 할 수 있다.

대해 설명할 수 있어야 하며, 그 내용에 대해 도덕적·법적 책임을 질 수 있어야 한다”는 기준을 명시하고 있으며, Gen AI는 자율적 판단 능력과 책임 주체로서의 요건을 충족하지 못하기 때문에 저자 자격을 부여할 수 없다는 것이 현재의 공통된 견해이다.⁹⁾

그럼에도 불구하고, 일각에서는 Gen AI가 연구 과정에 실질적으로 기여한 경우, 일정한 조건 하에서 공동저자적 지위를 고려할 수 있다는 주장도 제기되고 있다[17-20]. 예컨대 논문 초안 작성, 데이터 정리, 통계 처리 등에서 Gen AI의 기능이 유의미한 역할을 수행하였다면, 인간 연구자의 감독과 편집적 개입을 전제로 일정 수준의 기여를 인정할 여지가 있다는 것이다. 이러한 논의는 저자 자격을 전통적 책임의 기준에서만 판단하기보다는, 기여도 (contribution)에 따른 상대적 권한의 관점에서 재해석하려는 시도로 이해될 수 있다. 특히 Gen AI의 작동 과정에 대한 로그 기록과 결과물에 대한 역추적이 기술적으로 가능해질 경우, AI가 수행한 역할과 그 영향력을 보다 명확하게 규명할 수 있게 되며, 이에 따라 책임 분담의 제도화에 대한 논의가 한층 확대될 수 있다[21].

그러나 이 문제의 본질은 단지 AI를 저자로 등재할 수 있는가의 여부보다는, 연구 결과에 대한 최종 책임의 귀속과 설명 가능성의 확보에 있다. 따라서 학문 공동체는 Gen AI의 사용 여부와 그 기여 범위에 대한 명확한 정보 공개를 연구자에게 요구해야 하며, 연구자는 생성된 산출물 중 AI가 개입한 부분을 명시적으로 구분하고, 해당 내용에 대한 모든 설명 책임은 인간 저자에게 전적으로 귀속됨을 분명히 해야 한다. 특히 텍스트, 이미지, 통계자료 등에서 AI가 생성한 결과가 포함된 경우, 그 출처와 생성 방식에 대한 투명한 고지가 필수적이다.

결국 Gen AI의 연구 참여는 ‘공동저자’라는 법적 지위의 부여 여부에 앞서, ‘책임 있는 활용’과 ‘투명한 공표’라는 윤리적 기준에 따라 관리되어야 한다. 인공지능이 점차 학술 생태계의 필수 도구로 자리매김하고 있는 현재, 연구자는 저자의 개념과 책임의 범위를 재정립해야 하며, 학술기관과 출판사는 이에 대응하는 구체적인 정책과 기준을 제시할 의무가 있다. 이러한 제도적 정비와 윤리적 성찰 없이는, 학술 연구의 신뢰성과 정당성을 유지하는 기반 자체가 심각하게 훼손될 수 있다.

4. 연구자의 탈숙련화와 연구 역량 저하

Gen AI의 보편화는 연구자의 생산성과 효율성을 비약적으로 향상시킬 수 있는 기술적 도약을 가능케 했다. 반복적이고 규칙기반 작업의 자동화는 연구자의 시간과 노력을 절감하게 하며, 이로써 보다 창의적이고 고차원적인 사고에 집중할 수 있는 여지를 제공한다는 점에서 긍정적인 평가를 받을 수 있다. 그러나 이러한 기술적 진보는 동시에 연구자의 핵심 역량이 점진적으로 약화되는 ‘탈숙련화’ 현상을 유발함으로써, 연구윤리와 학문적 자기규율의 관점에서 중대한 문제를 제기한다[2].

예컨대 문헌 검색, 자료 요약, 번역, 초안 작성, 통계 분석 등 다양한 작업을 자동화함으로써

9) 국제학술지출판윤리위원회(Committee on Publication Ethics, COPE), 세계의학편집인협회(World Association of Medical Editors, WAME), 국제의학학술지편집인위원회(International Committee of Medical Journal Editors, ICMJE) 등 주요 국제 학술 출판기관들은 AI 도구가 연구 결과물의 정확성, 진실성, 독창성에 대한 책임을 질 수 없다는 점을 근거로 하여, AI를 논문의 저자 자격에서 명백히 배제하고 있다. 이들 기관은 AI 도구의 활용에 따른 모든 책임은 전적으로 인간 저자에게 귀속된다는 원칙을 명확히 하고 있으며, 특히 AI가 생성한 텍스트를 적절한 인용 없이 사용하는 행위는 표절로 간주될 수 있음을 강하게 경고하고 있다[11-13]. 미국의학협회저널네트워크(JAMA network) 또한 연구 진실성의 보장을 저자의 책임으로 규정하면서, AI 도구의 활용은 연구 설계 또는 방법론의 일부로 제한적으로 허용되어야 함을 권장하고 있다[16].

써 연구자의 업무 부담을 크게 경감시킬 수 있다는 점은 긍정적으로 평가될 수 있다. 그러나 동시에 자동화는 연구자가 연구 과정 전반에 능동적으로 참여할 기회를 제한하며, 결과물 중심의 비판적 검토 없는 수동적 연구 수행에 익숙해질 가능성 또한 배제할 수 없다. 이러한 위험은 특히 신진 연구자들에게 심각하게 작용할 수 있다[22]. 전통적으로 연구 역량은 연구 설계, 방법론적 고안, 문헌 해석, 개념 분석 등 다양한 학문적 실천을 반복적으로 수행하는 과정 속에서 점진적으로 형성되어 왔다. 그러나 최근에는 연구 초기 단계부터 Gen AI의 기능적 지원에 의존하는 사례가 증가하고 있으며[23,24], 이에 따라 기초적인 학문 훈련 없이 외형적으로 완성된 연구 산출물이 가능해지는 환경이 조성되고 있다.¹⁰⁾ 이러한 현상은 장기적으로 문제 해결 능력, 비판적 사고력, 학문적 자율성의 약화를 초래하며, 나아가 학문 공동체 전체의 질적 저하로 귀결될 수 있다.

더불어 Gen AI가 생성한 텍스트나 분석 결과는 문법적 완결성과 논리적 일관성을 갖춘 경우가 많기 때문에, 연구자가 이를 비판적으로 검토하지 않고 수용할 위험이 있다. 이와 같은 무비판적 수용은 연구 정당성에 대한 자기검열 기능을 약화시키고, 학문적 엄밀성을 저해하는 구조적 요인으로 작용할 수 있다.

따라서 탈숙련화 현상에 효과적으로 대응하기 위해서는 Gen AI를 단순한 결과물 생성의 도구로서 수동적으로 활용하는 데 그치지 않고, 오히려 비판적 사고를 촉진하는 지적 매개체로 적극적으로 활용하려는 태도 역시 필요하다. 즉 인공지능이 산출한 내용을 무비판적으로 수용하기보다는, 이를 면밀히 분석하고 자신의 언어로 재구성하며, 독립적 판단을 통해 연구자로서의 고유한 해석과 통찰을 형성하려는 노력이 요청된다. 아울러 교육기관과 연구기관은 신진 연구자 교육에 있어 단순한 기술 활용 역량뿐 아니라, 논리 구성력, 분석적 사고, 방법론적 정교성 등 핵심 학문 역량을 함양하는 방향으로 교육과 연구 훈련을 재설계할 필요가 있다.

결국 Gen AI로 인한 탈숙련화 문제는 단순히 개별 연구자의 능력 저하에 그치지 않으며, 장기적으로는 학술 생태계 전반의 지적 긴장감과 탐구 윤리성이 약화될 수 있는 구조적 문제로 이어질 수 있다. 이는 개인의 연구 방식에 국한되는 것이 아니라, 학문 공동체의 지속가능성과 연구 진실성을 보장하기 위해 사회적 관심과 제도적 대응이 요구되는 중대한 과제라고 할 수 있다.

III. Gen AI(Artificial Intelligence) 활용을 위한 연구윤리의 규범적 지향

앞서 검토한 윤리적 논의들은 Gen AI의 도입이 단순한 기술적 보조 수단의 채택에 머무르지 않고, 연구 수행 방식과 내용 구성은 물론 연구자에게 요구되는 진실성과 책임성에 구조적 변화를 야기하고 있음을 명확히 보여준다. 이러한 맥락에서 Gen AI 시대의 연구윤리는 기술 활용의 효율성이나 생산성에 관한 논의에 그칠 수 없으며, 연구 행위의 존재론적 지위와 학문 공동체 내에서의 의미를 재조명하는 철학적 성찰을 수반해야 한다. 본 장에서는 Gen AI의 학술적 활용에 있어 반드시 고려되어야 할 네 가지 규범적 지향, 즉 진실성과 투명성, 지식 재산

10) 한국연구재단은 2024년과 2025년에 걸쳐 Gen AI에 대한 연구자들의 인식 수준을 조사하고 그 결과를 보고서로 발간하였다. 이 두 보고서에 따르면, 2024년에는 Gen AI를 전혀 사용하지 않는다고 응답한 연구자의 비율이 2023년에 비해 현저히 감소한 것으로 나타났다. 예를 들어, 연구 설계 단계에서 Gen AI를 전혀 활용하지 않는다고 응답한 비율은 2023년 63.8%에서 2024년 52.8%로 크게 하락하였다. 이는 향후 다양한 형태의 Gen AI가 지속적으로 개발·보급됨에 따라, 해당 비율이 더욱 감소할 가능성을 시사한다[23, 24].

권의 존중, 분야별 규범의 정립, 그리고 전문적 역량의 고양을 제시하고자 한다.¹¹⁾

1. 진실성과 투명성

학술 연구는 언제나 진리 추구를 근간으로 하는 신뢰 기반 위에서 수행되어야 하며, 이는 인공지능의 활용 여부와 관계없이 변하지 않는 학문적 윤리의 기초이다. 특히 Gen AI를 연구의 과정이나 결과에 부분적으로 활용하는 경우, 연구자는 AI가 생성한 정보의 타당성과 정확성을 사전에 검토하고, 사후적으로도 그 결과물에 대한 철저한 검증을 수행할 책무를 가진다. 이는 단순히 산출물의 진위를 확인하는 절차를 넘어서, 해당 결과물이 어떠한 근거와 논리를 바탕으로 형성되었는지, 그리고 어떠한 맥락에서 사용되었는지를 투명하게 설명할 수 있어야 함을 의미한다.

Gen AI는 그 작동 방식의 특성상 허구적 문헌을 사실인 양 생성하거나, 기존 자료를 왜곡하여 사실처럼 제시할 수 있으며, 이러한 오류를 인식하지 못한 채 결과물을 연구에 수용할 경우, 학문적 신뢰성과 정당성이 심각하게 손상될 수 있다.¹²⁾ 실제로 존재하지 않는 문헌을 인용하거나, 정확하지 않은 데이터를 바탕으로 연구 성과를 발표하는 행위는 연구자의 개인 윤리 차원을 넘어, 학문 공동체 전체의 신뢰 기반을 약화시키는 구조적 위협으로 작용할 수 있다.

따라서 연구자는 Gen AI의 활용 전반에 걸쳐 결과물에 대한 실질적 책임을 분명히 인식해야 하며, AI 기술이 어떠한 방식으로 사용되었는지, 생성된 내용이 어떠한 기준을 바탕으로 선별되고 조정되었는지를 명확하게 기술해야 한다. 특히 인용된 정보의 출처를 투명하게 공개하고, AI 활용의 범위 및 의존 수준에 대해 구체적으로 서술하는 것은 연구윤리의 핵심 원칙으로 인식되어야 한다. 이러한 설명 가능성과 투명성은 단지 기술적 요구가 아니라, 디지털 전환 시대의 학술 생태계가 지속가능성과 공신력을 유지하기 위한 필수적인 윤리적 조건이라고 할 수 있다.

2. 지식 재산권 존중

Gen AI는 광범위한 데이터셋을 기반으로 학습하여 텍스트, 이미지, 그래프, 도표 등 다양한 형태의 콘텐츠를 생성한다. 이 과정에서 AI가 참조하거나 내재화한 자료에는 저작권으로 보호받는 원저작물이 포함될 수 있으며, 결과적으로 생성된 산출물이 기존 창작물과 구조나 내용, 표현 방식에 있어 유사성을 가질 경우, 연구자는 의도치 않은 지식 재산권 침해 혹은 표절 문제에 직면할 수 있다. 특히 AI가 만들어낸 결과물이 원본의 창의적 요소를 일정 수준 이

11) 필자는 Gen AI의 연구 활용에서 제기되는 윤리적 쟁점들이 연구 전주기(research lifecycle)의 각 단계와 긴밀히 연계된다고 앞서 판단했고(주석 2번 참조), 이에 상응하는 규범적 지향을 제시하면 다음과 같다. 먼저 연구 기획 및 설계 단계에서는 연구 참여자의 기여와 책임을 명확히 규정하는 문제가 핵심적으로 부각되며, 이에 따라 저자 자격과 책임의 논란은 학문 공동체의 합의를 전제로 한 '분야별 규범 정립'이라는 지향으로 연결된다. 이어서 데이터 확보 단계에서는 개인정보와 저작권 보호가 중요한 쟁점으로 제기되며, 이는 '지식 재산권 존중'이라는 규범적 지향을 통해 데이터 활용의 합법성과 정당성을 보장한다. 다음으로 데이터 활용 단계에서는 인공지능이 수집된 데이터를 처리·분석하는 과정에서 발생할 수 있는 환각과 정보 오류가 문제시되며, 이는 결과의 '진실성 확보와 투명성' 보장이라는 규범적 지향과 맞닿는다. 마지막으로 결과 도출 및 논문 작성 단계에서는 인공지능 의존이 연구자의 비판적 사고와 창의성을 약화시킬 수 있는 위험이 나타나며, 이에 대응하는 규범적 지향은 연구자의 '전문적 역량 고양'이라고 할 수 있다.

12) 이러한 산출물을 연구자가 비판적 검증 없이 활용할 경우, 존재하지 않는 데이터를 꾸며내는 위조나 실제 자료를 조작하는 변조와 동일한 결과를 초래할 수 있으며, 이는 연구윤리의 근간인 진실성과 책임성을 심각하게 위협한다.

상 재현하거나 변형한 경우, 그 활용은 법적·윤리적 책임을 수반할 수도 있다.

따라서 연구자는 Gen AI를 통해 생성된 모든 결과물에 대해 출처의 명시 여부, 인용의 적절성, 수정 또는 편집의 정도 등을 신중하게 판단해야 하며, 표면적인 인용 표시만으로는 연구윤리상 요구되는 책임을 다했다고 보기 어렵다. 생성된 콘텐츠가 기존 창작물에 대한 2차적 저작물로 간주될 수 있는지, 혹은 독창적 저작물로 인정될 수 있는지를 면밀히 검토하고, 이에 대한 자기 성찰적 태도를 견지해야 한다. 이는 단순한 표절 회피를 넘어, 창작물의 정당한 권리를 존중하는 비판적 윤리 감수성의 실천이다.

더 나아가 고려해야 할 것은 텍스트에만 한정된 문제가 아니라는 점이다. 이미지, 도표, 표, 시각적 시뮬레이션과 같은 비언어적 자료 또한 동일한 수준의 저작권적 고려 대상이 되며, AI에 의해 생성된 시각 콘텐츠 역시 원저작자의 권리를 침해할 소지를 내포할 수 있다. 이러한 윤리적 주의를 법적 의무의 차원을 넘어서, 타인의 창작물에 대한 존중과 학술적 성실성을 구현하는 실천적 자세이며, 궁극적으로는 연구자 개인뿐 아니라 학문 공동체 전체의 신뢰를 지속시키는 핵심적인 연구윤리 원칙으로 기능해야 할 것이다.

3. 분야별 규범의 정립

Gen AI의 활용 방식은 학문 영역에 따라 상이하게 나타나며, 모든 분야에 동일한 규범을 적용하는 것은 학술적 적합성 측면에서 바람직하지 않다. 예컨대, 자연과학 분야에서는 실험 설계의 자동화, 대규모 데이터 분석의 정밀화, 반복 작업의 효율화 등 기능 중심의 활용이 주로 이루어지는 반면, 인문학과 사회과학의 학문적 전통은 개념의 정교한 구성과 비판적 사유, 그리고 논증의 치밀함을 연구의 토대로 삼는다. 따라서 Gen AI의 개입이 이러한 학문적 본질에 어떠한 변화를 가져올 수 있는지에 대해 더욱 심화된 논의가 요청된다.

이러한 학문적 특성과 방법론의 차이를 고려할 때, Gen AI 활용에 대한 윤리 원칙은 획일적인 규제보다는 각 분야의 자율성과 고유한 학술적 전통을 존중하는 방식으로 구성되어야 한다. 이 점에서 연구윤리의 보편 원칙(진실성, 책임성, 투명성 등)은 모든 학문이 공유해야 할 최소 기준을 제공하지만, 각 분야의 특수한 연구 맥락과 방법론적 차이를 반영하기 위해서는 분야별 규범 정립이 필수적이다. 실제로 일부 저널과 학술기관은 AI 사용에 대한 투명성을 강조하며 구체적인 지침을 마련하고 있으며, 그 내용은 연구 목적, 도구의 기능, 기여 정도 등에 따라 상이한 기준을 제시하고 있다[25-27]. 특히 저자성에 대한 판단 기준, 인용의 방식, AI 활용의 목적 및 범위에 관한 명확한 기술은 학술 출판의 정당성과 신뢰를 유지하는 데 필수적인 기준으로 작용한다.

이러한 흐름 속에서 각 학문 분야의 전문 학회, 연구윤리위원회, 학술지 편집위원회 등은 연구자들의 자율적 판단과 공동체적 논의를 기반으로, Gen AI의 적절한 활용에 관한 지침을 체계적으로 마련해야 하며, 이는 투고 규정, 심사 절차, 학술 평가의 기준에 일관되게 반영될 필요가 있다. 나아가 이는 단지 절차적 요건을 충족하는 수준을 넘어, 연구자의 창의적 주체성과 윤리적 책임을 함께 존중하는 제도적 기틀로 작용한다.

궁극적으로 Gen AI 활용의 적합성과 윤리적 정당성은 기술 수용의 수준이 아니라, 해당 학문 공동체 내에서 형성된 합리적 합의와 판단에 기반할 때 더욱 공고해질 수 있으며, 이는 AI 시대의 학술 생태계가 지속가능한 신뢰 구조 위에서 작동하도록 하는 데 결정적인 역할을 할 것이다.

4. 전문적 역량 고양

Gen AI의 급속한 확산은 연구자에게 단순한 도구 운용 능력을 넘어, 연구 수행 전반에 걸친 역량 구조의 근본적 재편을 요구하고 있다. 과거 연구의 성숙도는 수작업에 기반한 반복적 기술의 숙련도나 정밀성에 의해 평가되었으나, 이제는 그러한 기준만으로 연구자의 전문성을 설명할 수 없다. 오히려 AI가 산출한 정보의 신뢰성과 적합성을 비판적으로 식별하고, 이를 연구 목적과 문제의식에 맞게 재구성하는 분석적 사고와 해석 능력이 연구자의 핵심 자질로 부상하고 있다. 이때 전문성은 단순히 새로운 기술을 활용하는 능력에 국한되지 않고, 자동 생성된 결과물의 한계를 인식하며 이를 학문적 맥락 속에서 독립적으로 재해석할 수 있는 통찰력과 판단력을 포괄한다.

특히 학술적 훈련이 충분히 축적되지 않은 초기 연구자의 경우, Gen AI가 제공하는 자동화된 편의성에 과도하게 의존할 위험이 존재한다. 연구 설계, 자료 해석, 논리적 서술과 같은 기초적 역량이 내면화되지 않은 상태에서 결과 중심의 산출물에 매몰될 경우, 학문적 성숙을 위한 필수 과정이 생략될 수 있으며, 장기적으로는 문제 제기 능력의 약화, 창의적 사유의 빈곤, 학문 공동체에 대한 기여 축소로 이어질 가능성이 크다. 이러한 현상은 단지 개인의 역량 저하에 그치지 않고, 학술 생태계 전반의 지속가능성을 위협하는 구조적 요인으로 작용할 수 있다.

이와 같은 위험을 예방하고 연구자의 총체적 전문성을 제고하기 위해, 교육기관과 연구기관은 단편적인 도구 사용법 교육을 넘어서는 포괄적 AI 리터러시 교육을 체계적으로 확립해야 한다. 이는 AI의 작동 원리와 알고리즘의 구조, 데이터 편향과 불투명성에 대한 이해를 포함할 뿐만 아니라, 자동 생성된 결과에 대한 윤리적·사회적 함의를 성찰하는 훈련까지 포괄해야 한다. 더 나아가 논리적 추론 능력, 정합적 논증 구성, 창의적 문제 설정, 의미의 다층적 해석과 같은 고차원적 학술 역량을 함께 함양할 수 있는 융합적 교육 프로그램이 요구된다.

결국 이러한 교육은 단순히 새로운 기술에 적응하기 위한 방편이 아니라, 연구자가 Gen AI 시대에도 독립적 학문 주체로서의 정체성을 유지하고, 창의적이면서도 책임 있는 연구 수행을 지속할 수 있도록 하는 기반으로 작동할 것이다. 이는 개별 연구자의 능력 강화 차원을 넘어, 학문 공동체 차원에서 연구윤리의 실천을 제도적으로 뒷받침하는 핵심 장치로 기능하며, 장기적으로는 AI 시대 학술 생태계의 신뢰성과 지속가능성을 지탱하는 결정적 토대가 될 것이다.

IV. 결론을 대신하여: Gen AI(Artificial Intelligence) 시대의 연구윤리 재정립을 위하여

Gen AI의 도입은 연구 환경에 전례 없는 기술적 편의와 창의적 가능성을 제공하는 동시에, 학문 공동체가 오랫동안 유지해 온 윤리적 틀을 근본적으로 재검토할 것을 요청하고 있다. 종래의 연구윤리는 연구자의 자율성과 정직성, 책임성, 그리고 연구 대상에 대한 존중을 핵심 원칙으로 삼아 왔으며, 이는 싱가포르 선언을 비롯한 주요 국제 학술기구의 윤리 강령을 통해 제도화되어 왔다[28]. 그러나 Gen AI의 활용은 연구의 경계와 주체성을 모호하게 하며, 인간 연구자와 비인간적 인공지능 간의 기여와 책임을 명확히 구분하기 어려운 새로운 지형을 만들어내고 있다.

www.kci.go.kr

이러한 전환은 단순한 기술적 변화에 그치지 않으며, 연구 행위의 존재론적 기반과 가치 체계에 대한 근본적인 성찰을 요구한다. 연구자의 정체성과 저자성의 의미, 창의성의 주체, 책임과 설명 가능성의 귀속 문제는 더 이상 고정된 전제 위에서 유지될 수 없다. Gen AI의 확산은 기존 학술 질서를 뒷받침해 온 여러 핵심 개념의 재구성을 요청하며, 이는 변화하는 학술 환경 속에서 연구윤리를 역동적으로 재편할 필요가 있음을 시사한다.

따라서 연구윤리는 기술 발전의 수용에 그치는 소극적 규범이 아니라, 디지털 전환 시대의 환경 변화에 민감하게 반응하며 지속적으로 갱신·정립되어야 하는 실천적이고 유기적인 윤리 체계로 정의되어야 한다. 이는 일회성의 선언이 아니라, 기술과 윤리, 제도의 상호작용 속에서 끊임없이 재구성되는 열린 과정으로, 다음과 같은 다층적 노력이 필요하다.

첫째, 연구자 개인은 Gen AI의 활용 여부와 그 범위를 자율적으로 결정할 권한을 가지되, 동시에 그러한 활용을 통해 산출된 연구 결과물에 대해서는 전적인 설명 가능성과 책임을 수용해야 한다. 이는 AI의 기여 정도와 무관하게, 연구자가 자신의 이름으로 발표한 학술 성과에 대해 학문 공동체와 사회 전반의 신뢰를 보장할 책임 있는 주체로 간주된다는 전제를 따른다. 따라서 Gen AI의 도입은 단순한 효율성의 문제가 아니라, 연구자의 자기 규율과 책임성을 확장하는 새로운 연구윤리적 과제로 이해되어야 한다.

이러한 과제는 몇 가지 실천적 책무로 구체화될 수 있다. 무엇보다 ‘공개(disclosure)’가 요구된다. 연구자는 논문이나 보고서에서 AI 활용 여부와 그 범위를 명확히 밝힘으로써 연구 과정의 투명성을 보장해야 한다. 또한 ‘검증(verification)’이 필요하다. AI가 생성한 인용이나 분석은 연구자가 직접 사실성과 타당성을 점검하여 환각과 표절의 위험을 예방해야 한다. 나아가 ‘책임(accountability)’을 명확히 해야 한다. 최종 산출물의 진실성과 학문적 정당성은 어디까지나 연구자 본인에게 귀속되며, 이는 인공지능에 전가될 수 없다.

결국 연구자는 Gen AI를 단순한 대체물이 아니라 인간적 창의성을 자극하고 확장하는 매개체로 활용해야 한다. 이때 연구자는 그 산출물이 지니는 사회적 파급과 윤리적 함의를 성찰할 의무를 가지며, 이를 통해 연구윤리의 재구성과 진화를 가능하게 하는 주체로 자리매김한다.

둘째, 학문 공동체는 각 연구 분야의 특수성과 방법론적 요구를 반영하여 Gen AI 활용에 관한 명확한 가이드라인과 출판 윤리를 확립해야 한다. 특히 AI의 기여 범위와 방식, 그리고 인용 및 출처 표기의 기준을 제도적으로 규정하는 작업은 필수적이다. 일부 학술지는 이미 이러한 기준을 선도적으로 마련하고 있으나, 이는 개별 기관 차원의 권고에 머물러서는 안 되며, 학문 전반에 걸친 제도화와 연구자 교육 및 실천을 통한 내면화가 병행될 때 비로소 실질적 효과를 거둘 수 있다. Gen AI의 책임 있는 활용을 위한 윤리적 기반은 이러한 규범적 정립과 문화적 수용의 상호작용 속에서 형성된다.


무엇보다 이러한 제도적 정비는 출판 생태계의 규범 확립을 통해 구체화될 수 있다. 논문 투고 과정에서 AI 활용 여부를 명시하도록 의무화하고, “활용 여부 체크리스트”와 같은 절차를 도입하는 것은 국제적 규범과의 일관성을 유지하고 출판 윤리를 강화하는 중요한 수단이 된다. 더불어 표절이나 환각, 부정확한 인용을 탐지하는 프로그램의 도입, 그리고 저자 확인 절차의 강화는 형식적 규제를 넘어 출판물의 질을 보증하는 핵심 장치로 작용할 수 있다.

따라서 학문 공동체가 마련해야 할 가이드라인은 제도적 장치와 기술적 검증 체계를 포괄하는 복합적 틀 속에서 실효성을 확보해야 한다. 나아가 이러한 규범은 외적 강제가 아니라 연구자 스스로 내면화하고 실천할 수 있는 문화적 기반 위에 정착될 때, 비로소 Gen AI의 책임 있는 활용을 가능케 하는 지속가능한 학문 생태계가 구축될 것이다.

셋째, 제도적 차원에서는 연구자의 윤리적 역량을 체계적으로 강화할 수 있는 교육 체계의

확립이 시급하다. 이는 단순히 AI 도구를 다루는 기술적 숙련을 넘어, AI의 작동 원리와 한계를 비판적으로 이해하고, 그 결과물을 학술적 맥락 속에서 창의적으로 재구성할 수 있는 능력의 함양을 의미한다. 이를 위해 기존의 연구윤리 및 방법론 교육을 확장하여, AI 리터러시, 데이터 윤리, 알고리즘의 투명성과 책임성, 디지털 비판성 등을 포괄하는 교육과정이 마련되어야 한다. 이러한 교육 체계는 연구자 개인의 윤리적 성숙을 촉진하는 동시에, 학문 공동체 전반에 책임 있는 AI 활용 문화를 정착시키는 기반이 된다.

나아가 제도적 차원은 교육을 넘어 국가적 정책과 국제적 규범의 정비와도 긴밀히 연결되어야 한다. 국가 차원에서는 일관된 가이드라인과 예방적 규제 체계를 통해 연구 과정에서 발생할 수 있는 위험을 미연에 차단할 필요가 있으며, 국제적 차원에서는 주요 규범과의 정합성을 유지함으로써 글로벌 학술 생태계에서의 신뢰를 확보해야 한다. 결국 제도적 교육 강화와 정책적 장치, 그리고 국제적 조율은 서로 긴밀히 보완되며 Gen AI 시대 연구윤리의 새로운 지평을 열어가는 핵심 기반을 이룬다. 이처럼 연구자 개인의 윤리적 자기규율, 학문 공동체의 규범적 정립, 그리고 제도·정책 차원의 체계적 장치가 상호 보완적으로 작동할 때, 비로소 Gen AI의 책임 있는 활용을 위한 지속가능한 연구윤리 생태계가 마련될 수 있다.

결론적으로, Gen AI 시대의 연구윤리는 기술 수용 여부에 대한 단순한 찬반 논의를 넘어선다. 그것은 책임 있는 활용과 연구자 정체성의 재정의라는 두 축 위에서 구성되는 새로운 윤리 체계에 대한 탐색이며, 이는 연구자 개인뿐만 아니라 학문 공동체 전체와 국가·국제적 제도가 공동으로 구축해 나가야 할 미래 지향적 과제라고 할 수 있다. 

REFERENCES

1. Panda S, Kaur N. Exploring the role of generative AI in academia: opportunities and challenges. *IP Indian J Lib Sci Inf Technol* 2024;9(1):12-23. <https://doi.org/10.18231/j.ijlsit.2024.003>
2. Bender EM, Gebru T, McMillan-Major A, et al. On the dangers of stochastic parrots: can language models be too big? In: *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. Toronto, ON; 2021. pp.610-623.
3. Institute of Information & communications Technology Planning & Evaluation (IITP). *ICT Brief* 2025;15 [Internet]. ICT; 2025 [cited 2025 Aug 9]. <https://www.itfind.or.kr/streamdocs/view/sd;streamdocsId=P-glMGhVHqpRrUaDQBg5cqRc9g90bk9Njc0fl37cSsc>
4. Ambrogi B. Not again! Two more cases, just this week, of hallucinated citations in court filings leading to sanctions [Internet]. *LawSites*; 2024 [cited 2025 Aug 9]. <https://www.lawnext.com/2024/02/not-again-two-more-cases-just-this-week-of-hallucinated-citations-in-court-filings-leading-to-sanctions.html>
5. Dahl M, Magesh V, Suzgun M, et al. Hallucinating law: legal mistakes with large language models are pervasive [Internet]. *HAI*; 2024 [cited 2025 Aug 9]. <https://hai.stanford.edu/news/hallucinating-law-legal-mistakes-large-language-models-are-pervasive>
6. Aljamaan F, Temsah MH, Altamimi I, et al. Reference hallucination score for medical artificial intelligence chatbots: development and usability study. *JMIR Med Inform* 2024;12(1):e54345. <https://doi.org/10.2196/54345>
7. Moore-Colyer R. AI hallucinates more frequently as it gets more advanced — is there any way to stop it from happening, and should we even try? [Internet]. *Sapien*; 2024 [cited 2025 Aug 9]. <https://www.livescience.com/technology/artificial-intelligence/ai-hallucinates-more->

- frequently-as-it-gets-more-advanced-is-there-any-way-to-stop-it-from-happening-and-should-we-even-try
8. Rubin LM. ChatGPT and research ethics. In: Lin X, Chan RY, Sharma S, et al., editors. ChatGPT and global higher education. Starscholas Network; 2024. pp.179-188.
 9. Beechamp TL, Childress JF. Principles of biomedical ethics. Oxford University Press; 2019.
 10. Yun YS. Technical consideration for generative AI and copyright infringement issues. *J Softw Assess Val* 2024;20(3):1-11.
<https://doi.org/10.29056/jsav.2024.09.01>
 11. Committee on Publication Ethics (COPE). Authorship and AI tools [Internet]. COPE; 2025 [cited 2025 Aug 9].
<https://publicationethics.org/guidance/cope-position/authorship-and-ai-tools>
 12. World Association of Medical Editors (WAME). WAME recommendations on ChatGPT and chatbots in relation to scholarly publications [Internet]. WAME; 2025 [cited 2025 Aug 9].
<https://wame.org/page3.php?id=110>
 13. International Committee of Medical Journal Editors (ICMJE). Recommendations for the conduct, reporting, editing, and publication of scholarly work in medical journals [Internet]. ICMJE; 2025 [cited 2025 Aug 9].
<https://www.icmje.org/icmje-recommendations.pdf>
 14. Council of Science Editors (CSE). Recommendations for promoting integrity in scientific journal publications [Internet]. CSE; 2025 [cited 2025 Aug 9].
<https://www.councilscienceeditors.org/recommendations-for-promoting-integrity-in-scientific-journal-publications>
 15. Nature portfolio. Artificial intelligence (AI) [Internet]. Nature portfolio; 2025 [cited 2025 Aug 9].
<https://www.nature.com/nature-portfolio/editorial-policies/ai>
 16. JAMA Network Open. Instructions for authors [Internet]. Jama Network; 2025 [cited 2025 Aug 9].
<https://jamanetwork.com/journals/jamanetworkopen/pages/instructions-for-authors>
 17. Kim S. Abbreviation for East-West philosophy studies. *Stud Philos East-West* 2024(112):303-327.
<https://doi.org/10.15841/kspew.112.202406.303>
 18. Chen Z, Chen C, Yang G, et al. Research integrity in the era of artificial intelligence: challenges and responses. *Medicine* 2024;103(27):e38811.
<https://doi.org/10.1097/MD.00000000000038811>
 19. Polonsky MJ, Rotman JD. Should artificial intelligent agents be your co-author? Arguments in favour, informed by ChatGPT. *Australas Mark J* 2023;31(2):91-96.
<https://doi.org/10.1177/14413582231167882>
 20. Saygın DA, Kabakçı ADA. The use of controlled artificial intelligence as a co-author in academic article writing. *Eur J Ther* 2023;29(4):990-991.
<https://doi.org/10.58600/eurjther1801>
 21. He J, Houde S, Weisz JD. Which contributions deserve credit? Perceptions of attribution in human-AI co-creation. In: CHI '25: Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems. Yokohama, Japan: PACIFICO Yokohama; 2025. pp.1-18.
 22. Mhlanga D. Generative AI for emerging researchers: the promises, ethics, and risks. *SSRN Electron J* 2024.
<http://doi.org/10.2139/ssrn.4737492>
 23. National Research Foundation of Korea (NRF). 2023 Survey on research ethics awareness among university faculty members. NRF; 2024.
 24. NRF. 2024 Survey on research ethics awareness among university faculty members. NRF; 2024.
 25. Thorp HH. ChatGPT is fun, but not an author. *Science* 2023;379(6630):313.
<http://doi.org/10.1126/science.adg7879>

26. Springer. Artificial intelligence (AI) [Internet]. Springer; 2025 [cited 2025 Aug 9].
<https://www.springer.com/gp/editorial-policies/artificial-intelligence—ai-/25428500?srsltid=AfmBOoryHPSfGTWPU1RFVa4g4oIJZk5bEnNBj5Ohk7KrYvR1wmYWFkzM>
27. Elsevier. Generative· AI policies for journals [Internet]. Elsevier; 2025 [cited 2025 Aug 9].
<https://www.elsevier.com/about/policies-and-standards/generative-ai-policies-for-journals>
28. The Office of Research Integrity (ORI). Singapore statement on research integrity [Internet]. ORI; 2025 [cited 2025 Aug 10].
<https://ori.hhs.gov/singapore-statement-research-integrity>