

생성형 AI 기반 도서관 서비스 환경에서의 데이터 편향성 탐색과 대응 전략*

Exploring Data Bias and Mitigation Strategies in Generative AI-Driven Library Service Environments

이 정 미 (Jeong-Mee Lee)**

초 록

본 연구는 인공지능(Artificial Intelligence, 이하 AI) 기술이 도서관 환경에 미치는 영향을 살펴보고 AI 시대에 도서관 정보서비스가 직면한 윤리적 도전과 과제(데이터 편향성 문제를 중심으로)를 심층 분석하고, 이를 극복하기 위한 도서관과 사서의 역할 및 도서관 정보서비스의 전략적 전환에 관한 제언을 제시하고자 방대한 문헌을 종합 분석하여 결과를 제시한 연구이다. 이를 위해 다음과 같은 세 가지 연구 질문을 제시하였다. 먼저 AI 기술이 도서관의 핵심 서비스를 어떻게 변화시키고, 이 변화 과정에서 발생하는 주요 윤리적 문제는 무엇인지 살펴보았다. 이후 AI로 인해 발생하는 데이터 편향과 같은 윤리적 문제점을 극복하기 위한 도서관과 사서의 전략적 역할에 대하여 분석하였다. 마지막으로 도서관 서비스 환경에서 AI 기술을 안정적으로 적용하기 위한 도서관과 정보 전문가의 전략 및 정책 방향을 제시하고자 하였다. 결론적으로 도서관은 AI 기술의 수동적 도입을 넘어 데이터 편향성과 같은 윤리적 문제를 해결하는 사회적 책무의 주체로 자리잡아야 하며 리터러시 교육의 다면화 속에 비판적 리터러시의 중요성이 강조되었다. 마지막으로 기술도입의 전 과정에서 명확한 가이드라인 수립과 알고리즘 감사 절차를 포함하는 윤리적 거버넌스 체계 마련을 핵심적으로 강조하였다.

ABSTRACT

This study examines the impact of Artificial Intelligence (AI) technologies on library environments and provides an in-depth analysis of the ethical challenges and issues—particularly data bias—that library information services face in the AI era. By synthesizing and critically reviewing a broad body of literature, the study offers insights into the roles that libraries and librarians must assume to address these challenges and presents recommendations for the strategic transformation of library information services. The research addresses three key questions: (1) how AI technologies are transforming core library services; (2) what ethical issues emerge throughout this transformation; and (3) how to present strategic and policy directions for the stable implementation of AI technologies in library service environments. Based on this analysis, the study proposes directions for the strategic role transition of libraries and information professionals to ensure the responsible and stable integration of AI technologies into library service environments. The findings emphasize that libraries must move beyond the passive adoption of AI and position themselves as responsible agents in addressing ethical issues arising from AI-driven systems. The study highlights the growing importance of critical literacy within increasingly diversified literacy education frameworks and underscores the need for a comprehensive ethical governance structure – including clear guidelines and algorithmic auditing procedures – across all stages of AI implementation.

키워드: 생성형 AI, 데이터 편향성, 정보서비스, 비판적 리터러시, AI 윤리
Generative AI, Data Bias, Information Services, Critical Literacy, AI ethics

* 이 논문은 서울여자대학교 학술연구비의 지원에 의한 것임(2025-0052).

** 서울여자대학교 사회과학대학 문헌정보학과 교수(jmlee@swu.ac.kr)

논문접수일자 : 2025년 11월 28일 논문심사일자 : 2025년 11월 29일 게재확정일자 : 2025년 12월 12일
한국비블리아학회지, 36(4): 183-205, 2025. <http://dx.doi.org/10.14699/kbiblia.2025.36.4.183>

© Copyright © 2025 Korean Biblia Society for Library and Information Science

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. 서론

1.1 연구의 필요성

빅데이터나 상황인식 컴퓨팅 확산 등과 같은 새로운 정보기술의 발달과 확산 시기에도 매번 도서관은 이러한 정보기술의 적절한 활용이 도서관 정보서비스에 있어 도전이며 기회임을 인식하고 이를 전략적으로 활용하기 위해 꾸준히 노력해왔다.

인공지능(Artificial Intelligence, 이하 AI)은 인간의 지각, 학습, 추론, 자연어 처리 능력을 컴퓨터로 구현하는 기술로서(김태영 외, 2021), 사회 전반의 구조적 변화를 이끄는 핵심 동력으로 자리 잡았다. ChatGPT나 Gemini 같은 생성형 AI의 자연어 생성·응답 제공 능력은 이 용자에게 보다 친화적인 학습 지원 같은 교육적 용도로의 활용에 용이하다는 잠재력을 나타내는 것이라 할 수 있으나 생성형 AI가 제시하는 응답이 완벽하게 신뢰할 수 있는가에 대해서는 상당수의 전문가가 문제를 제기하고 있는 현실이다(이정미, 2023). 그럼에도 불구하고 지금 우리 사회에서 AI는 단순한 기술적 도구를 넘어 지식 정보의 생성, 유통, 접근 방식을 근본적으로 재편하고 있으며, 이러한 패러다임의 전환은 도서관의 전통적 역할에 중대한 도전과 기회를 동시에 제공한다. 이미 많은 도서관은 챗봇과 디지털 어시스턴트를 통한 이용자 문의 응대, AI 기반의 체계적 문헌 고찰(living systematic literature reviews: SLRs) 지원, AI를 활용한 지식 발견(Knowledge Discovery) 등 AI 기술을 적극적으로 도입하고 있거나 도입을 고려하고 있기도 하다(Cox & Mazumdar, 2024).

이는 AI 기술의 도입이 더 이상 선택이 아닌 필수조건이 된 현실을 보여주는 것이라 할 수 있다. 역설적으로 바로 이런 이유 때문에 도서관의 역할은 그 어느 때보다 중요해졌다고도 할 수 있다.

기술 플랫폼과 미디어 등 기존 정보 게이트키퍼에 대한 대중의 신뢰가 약화되는 상황 속에서, 도서관은 사회적으로 가장 신뢰받는 기관 중 하나로서(BrodeFrank, 2025) 정보에 대한 신뢰성 약화에 대응할 독보적인 위치에 있다. 따라서 본 연구는 도서관이 AI 기술을 단순히 수용하고 소비하는 주체로서가 아니라, 알고리즘 시대의 윤리적 문제를 해결하는 사회의 필수적인 공공 기관으로서 역할을 재정립해야 한다는 문제의식에서 출발한다.

AI 기술의 혁신적 잠재력 이면에는 ‘알고리즘적 편향성(Algorithmic Bias)’이라는 심각한 윤리적 문제가 존재한다. AI 편향성이란 “AI 실행 주기에서 정확하지 않은 가설로 인해 근본적으로 불공정하고 차별적인 결과를 만들어 내게 되는 특성”으로 정의될 수 있다(Varsha et al., 2023). 이러한 편향성은 알고리즘이 학습 데이터에 내재된 사회적, 역사적 편견을 그대로 학습하고 증폭시키면서 불공정한 결과를 초래하는 방식으로 나타나게 된다. 예를 들면 미국 법원에서 피고인의 재범 가능성을 예측하기 위해 개발되었던 COMPAS 소프트웨어가 흑인들에게 불리한 편파적인 판정을 내린다거나, 산소포화도 측정 의료기가 흑인의 피부에 사용할 때는 정확도가 떨어지게 되는 등의 사례들은 AI 편향성이 사회 전반에 미치는 차별적 영향을 명백히 보여주는 것이라 할 수 있다(Varsha et al., 2023).

몇몇 정보학자들은 이러한 문제들을 도서관에서 결코 낮설지 않은 문제로 인식한다. 이들은 듀이 십진분류법(DDC)이나 미국 의회도서관 주제명표목표(LCSH)와 같은 전통적인 도서관 분류 체계 역시 특정 문화나 집단을 배제하고 왜곡해 온 편견의 역사를 가지고 있으며(Berendt et al., 2023; Smith, 2022), 이는 소수 집단에 대한 ‘표상적 피해(representational harms)’와 자원 분배의 불평등을 만들어내는 ‘분배적 해악(allocative harms)’을 초래해왔다고 주장하고 있다(Berendt et al., 2023). AI 기술은 이러한 기존의 편향성을 전혀 없는 규모로 자동화하고 재생산할 위험을 내포한다고 볼 수 있다. 나아가 데이터 편향의 근원은 단순히 기술적 오류에 해당하는 문제로 볼 수 없다. 이는 AI 모델의 학습을 위해 글로벌 사우스(Global South¹⁾)의 저임금 노동력이 착취되는 ‘데이터 식민주의(data colonialism)’ 현상(Arora et al., 2023)을 보여주는 것과 같으며 결국 알고리즘에 내재된 편향이 체계적인 글로벌 불평등의 산물임을 보여준다고도 할 수 있는 것이다.

더욱이, 과거 사서와 아키비스트가 전문적 윤리 강령에 기반하여 수행했던 ‘인간 게이트키퍼(human gatekeeper)’의 역할이 영리적 목적이나 문제 발생 소지가 있는 데이터에 의해 좌우되는 ‘알고리즘 게이트키퍼(algorithmic gatekeeper)’로 전환되면서(van Otterlo, 2018), 정보의 공정성과 다양성을 위한 전문적 개입이 배제될 위험도 커지고 있는 것이 현실이다. 이는 도서관이 지켜온 지적 자유와 정보 접근의 평등이

라는 핵심 가치를 심각하게 훼손할 수 있으며, 사서가 주도하는 AI 활용 등을 비롯한 외부의 윤리적 관리 감독이 그 어느 때보다 시급함을 시사하는 것으로 판단 가능하다.

지금까지 도서관 분야의 AI 관련 연구는 주로 기술의 기능적 도입 방안이나 일반적인 윤리 원칙 제시에 머물러 있었다. 하주혜 외(2025)가 AI 리터러시에 대한 연구는 점차 증가하고 있으나 대부분 기술의 활용에 대해 강조하거나, 윤리적인 고찰이라 하더라도 탐색적 수준에 머물러 있어, AI가 초래하는 윤리적 문제를 해결하기 위한 구체적인 실천적인 방안을 모색하는 연구는 부족한 실정이라고 지적한 것처럼 도서관 환경에서 ‘데이터 편향성’이라는 특정 윤리 문제에 초점을 맞춰 사서와 도서관의 구체적인 역할 전환을 제시하는 심층적인 연구는 매우 미흡하다.

따라서 본 연구는 AI 기술의 수동적 수용을 넘어, 도서관이 데이터 편향성과 같은 윤리적 문제를 해결하는 사회적 책무의 주체로서 역할을 재정립하고 이를 위한 전략적 방향을 모색해야 한다는 필요성에서 출발한다. 도서관은 단순히 기술을 도입해 활용하기만 하는 소모자가 아니라, 정보의 윤리적 유통을 책임지는 핵심 기관으로서 알고리즘의 공정성을 감시하고 비판적 AI 리터러시 교육을 선도해야 한다고 보는 것이다. 본 연구는 이러한 역할 전환을 위한 이론적 토대를 마련하고, 향후 도서관 현장의 실천적 변화를 분석하는 경험적 연구의 기틀을 다지고자 한다.

1) 글로벌사우스(Global South)는 주로 남반구에 위치한 120여 개발도상국을 묶어서 지칭하는 용어다. 글로벌 사우스는 최근 몇 년 사이 미·중 패권 경쟁, 기후변화, 코로나 팬데믹, 러시아-우크라이나 전쟁 등을 겪으며 독자적으로 세력화된 목소리를 내기 시작해 국제 정치·경제적으로 주목받고 있다. 출처: 연합뉴스(https://news.einfomax.co.kr)

이를 위해 최근까지 출판된 학술 연구 논문 중 “생성형 AI”, “AI”, “데이터 편향”, “AI 윤리”와 같은 키워드를 명시한 논문들을 추출하고 이 논문들의 내용을 분석하여 AI와 데이터 편향성, AI 윤리 등 주요 범주별로 나누고 내용 분석을 통해 종합적인 주제를 추출하는 방식으로 선행 연구를 종합분석하는 연구 방법을 취하여 결과를 제시하고자 하였다.

1.2 연구 목적 및 연구 질문

본 연구는 AI 기술이 도서관 환경에 미치는 영향을 분석하고자 하는 연구이다. 무엇보다 데이터 편향성 문제에 초점을 두어 도서관 환경에서 정보전문가와 사서가 직면한 윤리적 도전을 심층적으로 탐구하며, 이를 극복하기 위한 전략적 역할 전환과 그 방향을 제시하는 것을 목적으로 한다.

연구 목적 달성을 위해 다음과 같은 세 가지 연구 문제를 제시하여 살펴보고자 하였다. 각각의 연구 문제는 상호 유기적으로 연결되어 있으며, 연구의 전개 과정을 이끌어가는 지표가 될 것이다.

- 연구문제 1. AI 기술은 도서관의 핵심 정보서비스를 어떻게 변화시키고 있으며, 이 변화 과정에서 발생하는 데이터 편향성과 같은 주요 윤리적 문제점은 무엇인가?
- 연구문제 2. AI 기반 정보서비스에서 발생하는 데이터 편향성 문제를 극복하기 위한 도서관과 사서의 전략적 역할은 무엇인가?
- 연구문제 3. 도서관 서비스 환경에서 AI 기술의 윤리적 안정성을 확보하기 위해 수

립해야 할 구체적인 전략과 정책 방향은 무엇인가?

2. 이론적 배경

2.1 AI 기술과 도서관 환경의 변화

AI 기술은 더 이상 단순한 업무 자동화 도구로서만 활용되는 것이 아니라 이 사회 전반의 패러다임을 재편하는 가운데, 도서관 정보서비스 분야 역시 근본적인 변화를 이끌고 있다. 생성형 AI를 필두로 한 정보 기술의 혁신은 정보의 생성, 접근, 관리 방식을 발전시키면서 도서관 업무의 효율성을 높이고 나아가 이용자 경험을 재구성할 무한한 가능성을 제시하고 있다 (Bridges et al., 2025). 과거의 도서관이 물리적 공간과 소장 자료를 중심으로 운영되었다면, 현재의 도서관은 AI를 통해 정보에 대한 접근성을 극대화하는 동시에 이용자 개인에게 맞춤형 지능형 서비스를 제공하는 방향으로 나아가고 있다. 이러한 변화는 짧은 시간에 소장 자료 관리에서부터 이용자 서비스 방식까지 도서관의 모든 핵심 기능에 많은 영향을 미치면서 미래 정보서비스 방향 결정에 있어 중대한 전략적 의미를 가지게 되었다. 이 기술적 변화는 단순히 도서관 기존 업무의 효율성 증대 수준을 넘어, 도서관의 본질적인 기능과 사회적 역할에 대한 재정립을 요구하는 전략적 접점이라 할 수 있다. AI로 인한 기술적 변화가 도서관 업무에 있어서 정보 접근성과 이용자 경험을 재정의하고 있으며, 도서관을 정적인 정보보관소에서 동적인 지식 탐색 및 창출의 공간

으로 탈바꿈시키고 있기 때문이다.

2.1.1 서비스의 자동화 및 지능화

AI 기술은 도서관의 핵심 서비스에 통합되어 자동화와 지능화를 촉진하고 있다. 도서관 업무 서비스의 자동화, 지능화를 보여주는 사례는 다음 <표 1>과 같다.

이 사례들은 모두 도서관 서비스가 이용자의 요청에 따라 수동적으로 반응하던 것에서 벗어나, 데이터 분석과 자동화를 통해 이용자의 잠재적 요구를 예측하고 선제적으로 대응하는 능동적 패러다임으로 전환되고 있음을 보여주는 것이라 하겠다.

2.1.2 이용자 경험의 재구성

이러한 기술적 변화는 이용자의 정보 탐색 및 활용 경험을 근본적으로 재구성하고 있다. AI 기반 시스템은 24시간 접근이 가능한 챗봇, 이용자의 잠재적 정보 요구까지 예측할 수 있는 맞춤

형 추천 서비스 등을 통해 과거의 정적이고 수동적인 정보 제공 환경을 역동적이고 상호작용적인 정보 환경으로 전환시키고 있다(Boateng, 2025; Cox, 2022). 이용자는 더 이상 도서관 운영 시간에 상관없이 방대한 정보 속에서 자신의 필요에 맞는 자원을 보다 효율적으로 탐색, 활용할 수 있게 되었다.

도서관이라는 정보 공간에 대한 경험뿐 아니라 정보탐색과 활용, 정보서비스 등을 기반으로 하는 이용자의 정보접근성에 대한 경험까지도 기존의 경험에 완전히 다른, 보다 더 자유로운 정보 탐색 환경의 경험을 더함으로써 도서관이라는 물리적 공간에 대한 경험이 재구성된다는 것이다.

2.2 도서관 정보서비스와 AI - 도전과 과제

AI와 연관된 기술 발전의 이면에는 AI 시스템에 내재된 데이터 편향성으로 인해 정보의

<표 1> 도서관 업무 서비스의 자동화 및 지능화 사례
(곽우정, 노영희, 2021; Cox & Mazumdar, 2022에서 재구성)

서비스 영역	기술 및 시스템	주요 기능
참고 정보서비스	챗봇 기반 참고 정보서비스	• 딥러닝 및 자연어 처리(NLP) 기술을 활용 • 이용자의 질문 의도를 분석하여 24시간 자동으로 답변을 제공하는 형태로 진화 • 단순 반복적인 문의에 대한 사서의 업무 부담 감소 • 이용자에게는 시공간 제약 없는 정보 접근성을 보장
맞춤형/ 개인화된 정보 추천	빅데이터 기반 도서 추천 키오스크	• 이용자의 연령, 성별, 관심사 등 개인 데이터 기반 • 이용자의 과거 대출 이력과 검색 패턴을 분석 • 잠재적 관심사에 부합하는 정보를 선제적으로 제공
자동화된 출입 및 대출/ 반납 시스템	SKT의 'ALGO' 시스템	• ALGO: Automated Library GO • 안면인식 기술로 이용자 인증 • 천장에 설치된 RFID 안테나로 도서 태그를 결합해 사용자 인증과 대출·반납 과정을 완전 자동화
자동화된 장서 점검	장서 점검 로봇	• 장서 점검 로봇이 서가를 자율 주행하며 책등의 제목과 저자명을 스캔 • AI 알고리즘이 이를 데이터베이스와 대조하여 도서의 위치 오류를 실시간으로 파악 • 장서 점검 업무의 효율성을 획기적으로 개선

공정한 접근이라는 도서관의 핵심 가치를 위협할 수 있다는 심각한 윤리적 도전이 존재한다. AI가 내리는 결정은 결코 객관적이거나 가치중립적인 것이 아니며 AI 시스템의 의사결정은 학습 데이터에 절대적으로 의존하기 때문에, 데이터 자체에 내재된 사회·역사적 편견은 알고리즘을 통해 그대로 재현되거나 때로는 오히려 증폭될 수도 있다(Varsha et al., 2023). 이는 특정 집단에 대한 체계적인 정보 소외로 이어질 수 있으며, 지식의 보편적 접근을 추구하는 공익 기관으로서의 도서관에게는 도서관의 근본적인 사명을 위협하는 위험 요소로 작용한다.

2.2.1 알고리즘 편향성의 근원과 양상

AI 편향성은 단순한 기술적 오류가 아니라, 사회·역사적 편견이 데이터와 알고리즘 설계 과정에 복합적으로 내재되어 발생하는 문제이다. AI 편향성을 만들어내는 근원은 다각적으로 분석할 수 있다.

먼저 편향된 학습 데이터에 기인한다. 도서관이 보유한 방대한 서지 데이터는 AI 학습을 위한 중요한 자원이지만, 동시에 학습 결과의 편향을 만들어내는 주요 원천이 될 수 있다. 예를 들어, 듀이십진분류법(DDC)이나 미국 의회도서관 주제명표목표(LCSH)와 같은 전통적인 분류 체계는 서구 중심적, 식민주의적 관점이 반영되어 있으며, 특정 집단에 대한 편견을 담고 있다(Berendt et al., 2023; Smith, 2022). 이러한 분류체계로 구축된 데이터를 AI가 학습할 경우, 과거의 편견이 시스템에 그대로 들어오게 되고, 자동 추천 및 검색 결과를 통해 더욱 강화될 수 있게 된다. 이는 1930년대 흑인 사

서 도로시 포터(Dorothy Porter)가 흑인 시인의 작품을 인종에 기반한 '노예(slavery)' 항목이 아닌 작품의 주제 내용에 따라 재분류하며 기존의 편향에 저항했던 사례에서 볼 수 있듯이, 문헌정보학에서 오랫동안 다루어져 온 유서 깊은 과제라고 할 수 있다(Berendt et al., 2023).

두 번째로는 알고리즘 자체의 결함 및 인간의 개입으로 인해 만들어진다. 데이터뿐만 아니라 알고리즘 설계 및 데이터 레이블링 과정에서도 편향이 발생할 수 있다. 개발자가 무의식적으로 자신의 편견을 주입시키게 되거나 그 사회의 사회적 통념이 알고리즘의 구조나 가중치 설정에 영향을 미칠 수 있으며, 특정 결과가 더 선호되도록 시스템을 유도할 수도 있다(Carpenter, 2025). 결국 AI 시스템은 그것을 만든 인간과 사회의 편견을 그대로 반영할 수 밖에 없는 것이다.

2.2.2 데이터 편향성이 초래하는 차별과 표상적 피해

데이터 편향성은 실제 이용자에게 명확하게 보이지는 않더라도 사회 전반에 심각한 부정적 영향을 남길 수 있다.

먼저 '차별'의 문제이다. AI 알고리즘의 편향은 특정 인구 집단에 대한 체계적인 차별로 이어질 수 있다. 실제로 의료 분야에서는 흑인 환자들이 백인 환자들에 비해 의료 지원 대상에서 낮게 평가되었고(Obermeyer et al., 2019), 금융 분야에서는 여성의 주택담보대출 신청이 부당하게 거부되었으며(Bansal et al., 2023), 미국 사법 시스템에서 사용된 재범 예측 소프트웨어(COMPAS)는 흑인 피고인의 재범 가능성을 더 높게 예측하는 인종적 편향을 보였

다(Varsha et al., 2023). 이러한 사례들은 도서관 환경에서도 특정 인종, 성별, 사회·경제적 배경을 가진 이용자 집단이 정보 접근 기회에서 소외될 수 있는 실질적인 위험이 존재한다는 것을 시사하는 것이다.

다음은 '표상적 피해'의 문제이다. Noble(2018)은 검색 엔진이 '흑인 소녀'와 같은 검색어에 대해 성적으로 대상화된 결과를 노출하는 등 소수 집단에 대한 부정적인 고정관념을 강화하는 문제를 지적했으며 이를 '표상적 피해'로 정의했다(Carpenter, 2025; Noble, 2018). 이는 단순히 잘못된 정보를 제공한다는 차원을 넘어, 특정 집단의 사회적 정체성을 왜곡하고 모욕함으로써 사회적으로 문제를 발생시키는 행위이다. 또한 알고리즘은 유색인종 창작자의 콘텐츠를 억제하고 그들의 목소리를 주변화함으로써 정보 생태계의 다양성을 해칠 수 있다. 이러한 표상적 피해는 도서관의 검색 및 추천 시스템에서도 동일하게 발생하게 되기 때문에, 이용자의 인식에 부정적인 영향을 미치고 지적 자유를 침해할 수 있게 되는 것이다(Blodgett et al., 2020).

2.2.3 데이터 식민주의와 노동 착취

AI 윤리 문제는 알고리즘의 편향된 결과에만 국한되는 것은 아니다. 즉, 정보 검색과 결과들에만 한정되는 문제는 아니라는 것이다. 이는 AI 기술 개발의 근간을 이루는 데이터 수집 및 가공 과정에 내재된 글로벌 불평등 구조에서도 발생한다.

먼저 거대 기술 기업들이 주로 글로벌 사우스(Global South) 지역의 저임금 노동력의 착취에 기반한 데이터 수집과 이를 자신들의 이익을 위해 활용하는 과정에서 새로운 형태의 착

취 구조가 발생하게 되는 데이터 식민주의(Data Colonialism) 문제이다. 이는 과거 식민주의가 자원을 착취했던 것과 유사하게, 데이터를 통해 경제적·사회적 불평등을 심화시키는 '데이터 식민주의'로 비판받고 있다(Arora et al., 2023). 이러한 구조는 AI 기술의 결과로 만들어지는 혜택이 특정 지역과 기업에 집중되는 반면, 데이터 제공 지역은 소외되는 딜레마를 낳는다. 이러한 데이터 식민주의 구조는 AI 개발의 이면에 존재하는 숨겨진 인간 노동 착취 문제를 가능하게 하는 경제적, 지정학적 조건을 형성한다. 고도로 자동화된 것처럼 보이는 AI 기술의 내면에는 저임금·저숙련 노동자들의 '숨겨진 노동'이 존재한다. 대표적으로, OpenAI는 ChatGPT의 유해성을 줄이기 위해 케냐의 저임금 노동자들을 고용하여 폭력, 혐오 발언, 성적 학대와 같은 유해 콘텐츠에 레이블을 부착하는 작업을 수행하게 했다. 이들은 시간당 1.32달러에서 2달러 사이의 낮은 임금의 노동과 함께 유해 콘텐츠로 인한 심각한 정신적 외상에 노출되었다(Arora et al., 2023; BrodeFrank, 2025; Perrigo, 2023).

마지막으로 대규모 언어 모델 운영에 소비되는 막대한 양의 전력과 물은 심각한 환경적 영향을 야기한다(BrodeFrank, 2025; Li et al., 2023).

이처럼 AI가 제기하는 윤리적 문제는 기술적, 사회적, 경제적 차원을 종합적으로 포함하는 복합적인 성격을 지니기 때문에 이를 대응하는 데 있어도 신중한 접근이 필요하다.

2.3 AI 리터러시, 비판적 리터러시

선행 연구들은 AI 기술이 도서관 환경에 미

치는 긍정적 영향과 더불어, 알고리즘이 사회·문화적 편견을 답습하고 증폭시키는 등의 문제를 지속적으로 경고해왔다. 이에 따라, 단순히 AI 도구를 활용하는 능력을 넘어 새로운 차원의 리터러시를 요구하는 상황에 이르렀다. AI 리터러시 연구 초기 Long과 Magerko(2020)는 ‘AI 리터러시’를 AI 기술을 효과적으로 활용하고 비판적으로 평가하는 능력으로 정의했다(Long & Magerko, 2020). 그러나 최근 연구에서 다수의 학자들은 AI가 사회에 미치는 영향이 더욱 커짐에 따라 기능적으로 AI 기술을 활용하는 방법을 습득하는 것을 넘어서 기술의 윤리적, 사회적, 정치적 맥락을 성찰하는 ‘비판적 AI 리터러시’로의 개념 확장이 필수적이라 주장하고 있다(Bridges et al., 2025; Zhang et al., 2025). 도서관계 뿐 아니라 교육계에서도 AI의 작동 원리와 사회적 영향을 비판적으로 성찰하는 ‘비판적 AI 리터러시(Critical AI Literacy)’ 교육의 필요성이 새로운 과제로 부상하고 있다.

2.3.1 기술 활용을 넘어선 비판적 접근의 필요성

AI 시대에 도서관의 교육적 역할은 기존의 정보 활용 교육을 넘어서는 근본적인 전환을 요구받고 있다. 기술의 기능적 활용법을 넘어 그 이면에 숨겨진 사회적, 정치적, 윤리적 맥락을 성찰하도록 안내하는 ‘비판적 AI 리터러시(Critical AI Literacy, CAIL)’ 교육으로의 전환이 시급하다는 주장이 나오는 이유이기도 하다(Bridges et al., 2025; Zhang et al., 2025). 비판적 AI 리터러시에는 AI가 어떻게 편향을 학습하고 사회적 불평등을 강화할 수 있는지를 이해하는 능력을 포함한다. 이는 도서관이 단순히

최신 기술을 도입해 제공하는 수동적인 공간을 넘어, AI 기술의 윤리적 사용을 주도하고 정보 사회의 공정성을 지키는 사회적 책무의 주체로 자리매김하기 위한 핵심 과제이다.

단순히 ChatGPT와 같은 AI 도구의 사용법을 가르치는 ‘기능적 리터러시’는 AI가 제기하는 복합적인 문제를 해결하는 데 명백한 한계를 가진다. 비판적 AI 리터러시는 AI 기술이 사회, 문화, 경제, 정치에 미치는 복합적인 영향을 이해하고 그 윤리적 함의를 비판적으로 성찰할 수 있으며, 책임감 있는 시민의 자세로 AI를 활용할 수 있는 능력을 의미한다(Zhang et al., 2025).

AI 시스템의 작동 원리와 사회적 영향을 깊이 있게 성찰하는 ‘비판적 리터러시’가 필수적인 이유는 무엇보다 먼저 지금과 같은 정보 과잉 생산의 시대에, 콘텐츠 생성 AI의 정작속에 정보를 활용하는 이용자에게 무엇보다 중요한 역량은 진짜 같은 가짜에 대한 판별력이라고 할 수 있기 때문이다. 생성형 AI는 방대한 학습 데이터를 기반으로 문법적으로는 완벽하지만 사실이 아닌 ‘그럴듯한 가짜’ 또는 ‘진짜 같은 거짓말’을 생성할 수 있다. 이런 이유 때문에 AI가 생성한 결과물을 비판적 검증 없이 맹신하는 것은 매우 위험하며, 정보의 정확성을 판단하는 비판적 평가 능력이 필수적이다(이정미, 2023).

다음은 현재 우리 주변 이용자의 리터러시 수준에 대한 지나친 자신감에 있다. Montesi et al.(2025)의 연구에 따르면, 이용자들은 자신의 AI 리터러시 수준을 실제보다 높게 평가하는 경향을 보인다. 자신들의 AI 리터러시 수준을 과대평가하는 것이다. AI 도구의 사용자 인터

페이스가 친화적이기 때문에 자연스럽게 자신이 아주 숙련된 이용자인 듯한 착각을 불러일으킬 수 있다는 것이다. 이러한 인식과 실제 역량 간의 격차는 도서관이 체계적인 교육을 통해 개입해야 할 필요성을 명확히 보여주는 지점이라 하겠다(Montesi et al., 2025).

마지막으로 AI를 사회적, 정치적, 경제적 맥락과 분리된 기술적으로 중립적인 도구로만 제시하는 교육은 AI에 내재된 편향성과 불평등 문제를 간과하게 만들어 무비판적으로 AI에 의해 만들어진 가치관을 그대로 받아들이는 무조건적 정보 접근과 수용에 이르게 된다. 이러한 접근은 도서관과 정보학 분야가 추구하는 비판적 정보 리터러시와 사회 정의의 가치에 부합하지 않으며, AI의 공정성 문제를 해결하는 데 근본적인 한계를 나타내는 것이 아닐까 판단된다(Bridges et al., 2025).

2.3.2 비판적 AI 리터러시의 핵심 역량

선행연구에서 확인할 수 있는 비판적 AI 리터러시 교육 내용은 이용자에게 AI 시스템과 그것이 작동하는 사회·경제적 맥락, 그리고 시스템과 상호작용하는 자신을 평가할 수 있는 다층적인 비판적 시각을 제공해야 한다고 강조한다. 기존 연구들을 종합하면, 비판적 AI 리터러시는 시스템 및 데이터에 대한 책임, 사회·경제적 맥락에 대한 성찰, 이용자 중심의 비판적 실천이라는 세 가지를 핵심 역량으로 포함하고 있다.

시스템 및 데이터에 대한 책임은 기술 시스템 자체에 대한 투명성과 책임성을 요구하는 능력을 일컫는다. 이용자는 AI 시스템이 어떤 데이터로 학습되었는지 투명하게 공개하도록 요구

하고(Zhang et al., 2025), AI의 의사결정 과정이 ‘블랙박스’처럼 불투명하게 유지되어서는 안 되며 그 작동 원리와 판단 근거를 설명할 수 있어야 한다. 이는 ‘투명성’과 ‘설명가능성’으로 인식되고 있다(Ntoutsis et al., 2020).

사회·경제적 맥락에 대한 성찰은 AI 기술 개발과 배포의 이면에 숨겨진 전 과정을 이해하는 것을 의미한다. AI 모델 학습에 막대한 에너지와 자원이 소모되는 환경적 영향과, 데이터 레이블링 등과 AI 개발 과정에 내재된 저임금 ‘숨겨진 노동’ 착취와 같은 사회·경제적 문제의 맥락까지도 비판적으로 인식하는 능력이 요구된다(Zhang et al., 2025).

마지막으로 이용자 중심의 비판적 실천은 이용자가 AI와 책임감을 가지고 상호작용하기 위해 개발해야 할 필수적이고 능동적인 기술에 초점을 맞춘다. AI가 제시하는 결과물을 무비판적으로 수용하려는 ‘자동화 편향(automation bias)’의 위험성을 인지하고, AI가 생성한 정보의 정확성, 신뢰성, 편향성을 항상 비판적으로 검증하는 태도를 훈련하는 것이 핵심이다(Carpenter, 2025; Zhang et al., 2025).

3. AI 시대 도서관 정보서비스의 전략적 전환

AI 기술이 사회 전반의 구조를 재편하면서 인류의 지식 보고인 도서관 역시 전례없는 변화의 기로에 서 있다고 할 수 있다. 단순히 기술 도입을 이야기하는 차원을 넘어 도서관의 핵심 기능과 정보서비스의 본질에 대한 근본적인 질문들이 파생되는 것도 지금의 현실이다. 이러

한 현실속에서 많은 학자들이 기술 변화에의 적응뿐 아니라 AI 알고리즘에 대응하는 도서관의 전략적 전환을 고민하고 있다. 이에 도서관 조직 및 정보서비스 전략 수립을 위한 시작점으로 지금까지의 다양한 학자들의 연구를 종합 분석하여, AI 기반 정보서비스의 가능성, 데이터 편향성과 같은 AI 알고리즘에 대응하는 도서관의 책무, 도서관 리더러십 교육의 개념 재정립, 정보전문가의 역량 및 조직 전략이라는 네 개 주제의 분석을 통해 도서관 정보서비스의 전략적 전환을 위한 제언을 도출하고자 한다.

3.1 AI 기반 정보서비스의 잠재력에 대한 탐색과 적용

3.1.1 핵심 서비스의 전환: AI 도입 잠재력 분석

AI 기술은 도서관의 전통적인 서비스를 다각도로 혁신하며 새로운 가치를 창출하고 있다. 국내외 도서관과 관련 기술 기업들은 이미 다양한 AI 기반 서비스를 도서관에 도입하며 정보자원 관리 및 접근성 향상, 이용자 지원 및 상호작용 강화, 학술 연구 지원의 고도화 등을 가능하게 함으로써 정보서비스 업무에 그 가능성을 입증하고 있다.

1) 정보 자원 관리 및 접근성 향상

AI 도서 추천 키오스크나 로봇 같은 AI 기술 상품은 방대한 장서 관리의 효율성을 높이고 이용자의 정보 접근성을 크게 개선한다. 빅데이터를 기반으로 이용자의 연령, 성별, 관심사 등을 분석하여 맞춤형 도서를 추천하는 AI

도서 추천 키오스크는 개인화된 정보 발견 경험을 제공한다(김태영 외, 2021). 또한, 여수 이순신도서관에 도입된 자동 도서 점검 로봇의 경우는 AI 알고리즘을 통해 수만 권의 장서를 단 몇 시간 만에 점검하고 잘못된 위치에 배열된 도서를 식별하여 사서의 장서 관리 부담을 획기적으로 줄여준다(김태영 외, 2021). 이는 사서가 단순 업무에 소비되는 시간을 줄여주어 보다 전문적인 업무에 집중할 수 있도록 환경을 조성하는 동시에, 이용자가 원하는 자료를 더 빠르고 정확하게 찾을 수 있도록 도와준다.

2) 이용자 지원 및 상호작용 강화

챗봇과 인공지능 기반 가상 비서는 시공간의 제약 없이 이용자 문의에 대응하며 상호작용을 강화할 수 있다. SK텔레콤이 개발한 지능형 가상 비서 'ALGO'는 안면 인식과 RFID 기술을 결합하여 도서관 출입과 도서 대출·반납을 자동화함으로써 이용자 편의를 극대화했다고 평가된다(김태영 외, 2021). 이와 같은 AI 챗봇은 24시간 내내 단순하고 반복적인 문의에 자동으로 응답하여 사서의 업무 부담을 줄이면서 이용자들은 언제든지 필요한 정보를 얻을 수 있어 서비스 만족도를 향상시킬 수 있는 기술적 지원을 가능하게 하였다.

3) 학술 연구 지원의 고도화

AI는 연구자들에게 정보 수집 및 분석 과정을 혁신할 수 있는 강력한 도구로 기능할 수 있다. 특히, 체계적 문헌 고찰(SLRs) 과정에서 AI를 활용하게 되면 방대한 양의 학술 문헌을 자동으로 점검하고 분석할 수 있기 때문에 연구자에게는 연구 시간을 단축하고 정확도를 높일

수 있는 도움을 주게 된다(Cox & Mazumdar, 2023). 더 나아가 AI 기반 지식 발견 도구는 기존에 발견하지 못했던 문헌 간의 연관성을 파악할 수도 있게 되고 이러한 새로운 발견을 통해 새로운 연구 가설을 제시함으로써 연구의 지평을 넓히는 데 기여할 수도 있다.

3.1.2 발생 가능한 윤리적 쟁점

AI 서비스가 제공하는 혁신적 이용 편리성의 이면에는 도서관이 반드시 해결해야 할 심각한 윤리적 문제들이 존재한다. 이러한 문제들은 도서관의 공공성과 신뢰성에 직접적인 영향을 미칠 수 있다.

1) 데이터의 프라이버시 및 감시

AI 시스템은 개인화된 서비스를 제공하기 위해 대량의 이용자 데이터를 수집하고 분석한다(Islam et al., 2025). 이 과정에서 이용자의 독서 이력, 검색 기록, 심지어 도서관 내 물리적 이동 경로까지의 민감한 개인정보가 무분별하게 수집될 위험이 존재한다. 이는 이용자의 사생활을 심각하게 침해할 수 있으며, 도서관이 이용자를 감시하는 주체가 되는 것처럼 인식될 수 있다는 우려를 낳는다(van Otterlo, 2018). 도서관은 이용자 데이터를 보호할 윤리적인 책무가 있으며, 데이터 수집 및 활용에 대한 명확하고 투명한 정책을 수립해야 한다. 우리나라의 경우 개인정보 보호법과 연계해 이용자 데이터 보호가 중요한 문제로 인식되고 있는 상황이다.

2) 투명성 및 설명 가능성 부족

많은 AI 알고리즘은 ‘블랙박스’처럼 작동하

여 어떠한 근거로 특정 도서를 추천하거나 정보를 필터링하는 것인지 그 의사결정 과정을 이해하기 어렵다(Ntoutsi et al., 2020). 이러한 투명성과 설명가능성의 부재는 알고리즘이 편향된 결과를 도출하더라도 이용자나 사서가 이를 인지하고 문제 발생시 적시에 문제를 시정하기 어렵게 한다(Ngulube, 2025). 이용자에게는 정보가 추천되는 근거를 알 권리가 있으며, 사서에게는 시스템의 작동 원리를 이해하고 그 공정성을 감독할 수 있는 자유가 있어야 도서관 서비스의 신뢰를 유지할 수 있을 것이다.

3) 지적 재산권 및 저작권 문제

ChatGPT나 Gemini 같은 생성형 AI의 대중화 이후 AI가 콘텐츠를 생성하고 요약하며 배포하는 과정에서의 저작권 침해 문제가 많은 관심의 중심에 있다. 예를 들어, 생성형 AI가 학습 데이터로 학습한 저작물과 비슷한 결과물을 생성하거나, 허가 없이 디지털 콘텐츠를 복제하고 배포하는 경우 법적 분쟁으로 이어질 수도 있다(Islam et al., 2025). 도서관은 AI를 활용한 서비스 제공 시, 저작권법을 철저히 준수하고 창작자의 권리를 보호해야 할 법적, 윤리적 책임을 확실하게 규정해야 한다.

이처럼 AI는 도서관 서비스에 전례 없는 혁신을 가져올 잠재력을 지니고 있지만, 데이터 프라이버시, 투명성, 저작권 등 복합적인 윤리적 과제를 동시에 안고 있다. 이와 함께 사회적 약자와 소수자를 소외시키고 정보 불평등을 심화시킬 수 있는 ‘데이터 편향’ 문제는 도서관의 핵심 가치와 직결되는 가장 시급하고 중요한 과제이다.

3.2 AI 알고리즘, 데이터 편향에 대응하는 도서관의 책무

3.2.1 알고리즘 편향의 원인과 유형

알고리즘 편향은 복합적인 원인에 의해 발생하며, 그 유형 또한 다양하다. 근본적인 원인을 이해하는 것은 효과적인 대응 전략을 수립하기 위한 첫걸음이다.

1) 데이터에 내재된 편향

가장 근본적인 원인은 학습 데이터 자체에 내재된 편견이다(Berendt et al., 2023). 데이터는 현실 세계의 복제품이라고도 할 수 있기 때문에 역사적·사회적으로 축적된 차별과 불평등을 그대로 반영하게 된다(Ntoutsi et al., 2020). 과거에 수집, 학습한 데이터에 특정 인종이나 성별에 대한 부정적 고정관념이 포함되어 있다면, 이를 학습한 AI 모델은 편향된 예측을 반복하는 결과를 갖게 된다. 또한, 특정 그룹의 데이터가 과소/과대 표집되는 표본 편향(sampling bias) 역시 심각한 문제다. 소수자나 비주류 집단의 데이터가 부족할 경우, AI는 이들을 제대로 인식하거나 공정하게 평가할 수 없기 때문이다(Berendt et al., 2023).

2) 알고리즘 시스템의 편향

사피야 우모자 노블(Safiya Umoja Noble)의 연구는 검색 엔진과 같은 정보 시스템이 구조적으로 특정 인종이나 젠더 그룹을 왜곡하고 소외시키는 방식을 폭로하고 있다(Carpenter, 2025; Noble, 2018). 이러한 '표상적 피해'는 알고리즘이 특정 집단을 부정적이거나 모욕적인 방식으로 묘사함으로써 발생하게 되고, 이

는 사회적 편견을 더욱 공고히 하는 결과로 이어지게 된다. 이는 개별 데이터의 문제를 넘어, 알고리즘을 설계하고 운영하는 시스템 전체에 편향이 깊숙이 내재되어 있음을 보여주는 것이라 할 수 있다. 사법 시스템 예측 소프트웨어에서 확인된 인종 편향의 오류나 금융서비스, 의료서비스 등에서 확인된 성차별, 인종 차별 사례 등이 알고리즘 시스템의 편향을 보여주는 사례들이다(Bansal et al., 2023; Obermeyer et al., 2019; Varsha et al., 2023).

3.2.2 편향성에 대한 도서관의 전통적 대응과 그 한계

정보 분류 체계에 내재된 편향성 문제는 도서관 역사에서 오랫동안 다루어져 온 주제이다. 듀이 십진분류법(DDC)이나 미국 의회도서관 주제명표목표(LCSH)와 같은 전통적인 분류 체계는 서구 중심적, 남성 중심적, 식민주의적 시각을 반영하고 있다는 비판을 받아왔다(Berendt et al., 2023). 이에 맞서 사서들은 부적절한 정보 분류 및 목록에 대응하며 정보 정의를 실현하기 위해 노력해왔다. 대표적인 사례 중 하나로 1930년대 하워드 대학교의 사서 도로시 포터(Dorothy Porter)는 당시 대부분의 백인 중심 도서관에서 만연해 있던 흑인 시인의 시집을 인종 차별의 색채가 가미된 '노예제'나 '식민' 관련 번호로 분류하는 관행을 거부하고, 저자의 인종이 아닌 작품의 내용에 따라 자료를 재분류함으로써 흑인 저자들의 지적 성취를 정당하게 평가하고 가시화했다(Berendt et al., 2023). 이처럼 도서관은 정보조직, 분류 체계의 편향에 맞서 싸워온 역사적, 철학적 기반을 가지고 있으며, 이는 AI 시대의 데이터 편향 문제에

대응할 수 있는 중요한 자산이 될 것이라 믿는다.

3.2.3 데이터 편향 극복을 위한 도서관의 전략적 역할

다양한 학자들이 도서관과 사서가 AI 데이터 편향 문제 해결을 위해 수행할 수 있는 구체적인 전략적인 역할을 제시하고 있다. 이 중 대표적인 논의는 감사나 완화의 주체 역할, 다양성을 증진하는 활동 강화, 참여적 메타데이터 생성으로 기존의 편향성을 희석해야 한다는 주장이다.

1) 편향성 감사 및 완화의 주체 역할

Smith(2022)는 도서관은 도입하려는 AI 시스템을 무비판적으로 수용해서는 안 되기 때문에 사서는 AI 알고리즘과 학습 데이터를 비판적으로 검토하고, 잠재적 편향성을 체계적으로 평가하는 편향성 감사(bias audit)의 주체가 되어야 한다고 주장하였다. 나아가 기술 전문가들과 협력하여 편향성을 완화하는 기술을 적용하고, 시스템이 도서관의 윤리적 기준을 충족하는지 지속적으로 감독하는 역할을 수행해야 한다는 주장이다.

2) 다양성 증진을 통한 대응

데이터 편향을 완화하는 가장 근본적인 방법 중 하나는 데이터의 다양성을 확보하는 것이다. 편향은 지나치게 하나의 그룹에 데이터가 집중되기 때문에 생기는 문제이기 때문이다. Berendt et al.(2023)은 다양성의 세 가지 구성 요소 개념을 활용하여 전략적으로 편향성에 대응할 수 있다고 주장한다. 이 구성 요소 개념은 다양성

(variety), 균형성(balance), 격차(disparity)를 의미한다. “다양성”은 얼마나 많은 다른 옵션이 있는지를 의미하는 것이고 “균형성”은 서로 다른 옵션들이 어떻게 분포되어 있는지를 나타낸다. 셋째, “격차”는 옵션들이 서로 얼마나 다른지를 측정하는 것이다. Berendt et al.(2023)은 도서관이 이 개념들을 적용하여 역사적으로 소외되어 왔거나 기록되지 않았던 지역 공동체 또는 소수자 그룹의 데이터를 의도적으로 수집하고 디지털화하여 보이게 함으로써 데이터의 다양성과 균형성을 높인다면 전체적인 데이터 편향을 줄일 수 있다 주장한다. 이는 AI가 이용자가 묻는 질문에 보다 포용적이고 공정한 결정을 내릴 수 있도록 학습에서부터 편향을 줄일 수 있게 하는 방법적인 측면에서 효과가 있을 것으로 판단된다.

3) 참여적 메타데이터 생성

BrodeFrank(2025)는 이용자가 참여한 분류 체계, 즉 이용자 스스로 만들어가는 크라우드소싱(crowdsourcing)과 같은 이용자 참여형 프로젝트를 제안했는데 이러한 분류 체계는 기존 분류 체계의 한계를 보완하는 효과적인 대안이 될 수 있다고 하였다(BrodeFrank, 2025). 도서관은 이용자들이 직접 소장 자료에 태그를 달거나 설명을 추가하는 참여적 메타데이터 생성 프로젝트를 기획함으로써, 단일한 시각에서 벗어난 다양한 관점의 데이터를 구축할 수 있게 된다. BrodeFrank(2025)는 이렇게 축적된 풍부하고 다각적인 메타데이터는 기존 분류 체계의 편향을 보완하고, AI 학습 데이터의 질을 향상시키는 데 중요한 기여를 할 수 있을 것이라 주장한다.

도서관이 데이터 편향 문제에 대응하는 사회적 책무의 주체로서 그 역할을 적절히 수행하기 위해서는 무엇보다 사서의 전문성 강화와 이용자의 비판적 역량 함양이 필수적이다. 이는 곧 AI 리터러시뿐 아니라 도서관이 지금까지 진행해 온 다른 모든 리터러시 교육의 패러다임이 근본적으로 전환되어야 함을 의미하는 것으로도 볼 수 있다. 다음은 AI 시대에 요구되는 새로운 리터러시 교육의 개념을 재정립하고, 그 구체적인 방향을 논의하는 부분이다.

3.3 도서관 리터러시 교육의 개념 재정립

이는 AI 시대 모든 리터러시 교육에 해당하는 문제일 수 있지만 지금 이 연구에서는 AI 관련 문제들을 정리하고 있으므로 이 부분에는 AI 리터러시라는 주제로 논의를 제한한다.

3.3.1 AI 리터러시의 핵심 구성 요소

AI 리터러시는 단일한 기술 활용 능력을 의미하는 것이 아니라, 다차원적 역량의 집합체이다. 여러 연구를 종합해 볼 때, AI 리터러시는 다차원적 역량이며 책임감 있는 AI 시민성을 담은 역량으로 강조되고 있다.

1) 다차원적 역량으로서의 AI 리터러시

AI 리터러시는 AI 기술의 작동 원리에 대한 기본적인 이해, AI 도구를 목적에 맞게 활용하고 적용하는 능력, AI가 생성한 결과물을 평가하고 새로운 가치를 생성하는 능력, 그리고 AI 기술 사용에 따르는 윤리적 쟁점을 인식하고 책임감 있게 행동하는 능력을 포괄한다(Hervieux & Wheatley, 2024). 즉 AI 리터러시가 추구하는 역량은 기술적 능력과 비판적 사고, 윤리적 성찰이 결합된 통합적 역량으로 이해되어야 한다(Zhang et al., 2025).

2) 책임감 있는 AI 시민성(AI Citizenship)

AI 리터러시 교육의 궁극적인 목표는 단순히 기술을 잘 사용하는 사용자를 만드는 것이 아니라 책임감 있는 AI 시민을 양성하는 데 있다. Hossain(2025)이 제시한 'AI 시민성 프레임워크(AI Citizenship Framework)'는 이를 위한 단계별 역량을 제시하고 있다(Hossain, 2025, <표 2> 참조).

3.3.2 전통적 리터러시에서 비판적 AI 리터러시로의 확장

AI 시대의 도서관 리터러시 교육의 핵심은 AI가 생성한 방대한 정보를 어떻게 비판적으

<표 2> AI 시민성 프레임워크-단계별 역량(Hossain, 2025)

단계	설명
AI 인식(AI awareness)	일상 속 AI의 존재를 인지하는 단계
AI 친숙성(AI familiarity)	AI 도구 사용법에 익숙해지는 단계
AI 리터러시(AI literacy)	AI의 원리를 이해하고 활용하는 단계
비판적·윤리적 AI 리터러시 (Critical and ethical AI literacy)	AI의 사회적 영향을 비판적으로 성찰하는 단계
AI 시민성(AI citizenship)	사회 구성원으로서 책임감 있게 AI 사용을 옹호하고 참여하는 단계

로 평가하고, 그 결과를 윤리적 책임감을 가지고 활용할 수 있는지에 대한 역량을 기르는 방향으로 재정립되어야 한다.

미국 학술연구도서관협회(ACRL)가 제시한 정보 리터러시 프레임워크는 AI라는 새로운 정보 환경에서도 여전히 유효하며, 새로운 맥락에 맞게 확장하여 적용 가능한 것으로 판단된다(James & Filgo, 2023).

좀 더 구체적으로 살펴보면 ACRL 프레임워크는 유연한 구성을 가지고 있기 때문에 사서들이 ChatGPT와 같은 새로운 기술을 이 프레임워크의 교육 과정에 더 쉽게 통합할 수 있도록 한다.

James와 Filgo는 ChatGPT(나아가 다른 생성형 AI들)가 ACRL 프레임워크의 여섯 가지 프레임 중 다수와 연결가능하며, 사서들이 왜 정보 리터러시 교육을 제공하는지 간결하게 제시하면서 전통적 리터러시 프레임워크가 AI 리터러시와도 밀접한 연관성이 있음을 주장하였다(James & Filgo, 2023).

Bridges et al.(2025)은 현재 문헌정보학(LIS) 교육 커리큘럼이 AI를 종종 기술적으로 중립적인 도구로만 다루는 경향이 있음을 비판하면서,

도서관 리터러시 교육을 기술의 사회적·정치적·경제적 맥락을 심도 있게 분석할 수 있게 하는 ‘비판적 AI 리터러시(Critical AI Literacy)’ 교육으로 전환해야 한다고 주장하기도 했다(Bridges et al., 2025). 이는 사서들이 AI 시스템의 편향성, 데이터 식민주의, 숨겨진 노동과 같은 문제를 이해하고, 이용자들이 이러한 문제에 대해 비판적으로 사고할 수 있도록 교육하는 것을 목표로 하는 것이라 하겠다.

지금의 정보 환경은 기존의 데이터 및 정보 리터러시 교육이 한 단계 더 진화해야 할 필요성을 제기한다. “진짜 같은 가짜” 정보의 등장과 이를 검증하는 비판적 사고 능력의 중요성이 그 어느 때보다 부각되는 것이 바로 지금의 AI 시대인 것이다.

이러한 문제에 대응하기 위해 몇몇 학자들은 비판적 AI 리터러시(Critical AI Literacy, CAIL)라는 새로운 개념을 제기하였으며 Zhang et al.(2025)은 CAIL을 위해 이용자가 AI 생성 결과물을 평가할 수 있는 구체적인 기준으로 RACBAC 기준을 제안(〈표 3〉)하고 있다(Zhang et al., 2025).

RACBAC 기준은 이용자가 불투명한 AI 시

〈표 3〉 RACBAC 기준 - AI 생성 결과물 평가를 위한 구체적인 기준(Zhang et al., 2025)

평가 기준	설명
관련성(Relevance)	생성된 결과물이 나의 연구 질문의 내용 및 깊이와 얼마나 부합하는가?
정확성(Accuracy)	생성된 콘텐츠의 사실관계가 신뢰할 수 있는 출처를 통해 검증 가능한가? 허위 정보나 조작된 내용은 없는가?
포괄성(Coverage)	결과물이 주제를 편향 없이 종합적으로 다루는가? 특정 관점이나 중요한 정보가 누락되지는 않았는가?
편향성(Bias)	결과물에 특정 이념, 문화, 인구통계학적 편견이 내재되어 있지는 않은가?
권위성(Authority)	결과물이 신뢰할 수 있는 출처를 기반으로 하는가? 제시된 출처 정보가 명확하고 검증 가능한가?
최신성(Currency)	정보가 해당 분야의 가장 최신 동향과 연구 결과를 반영하고 있는가?

시스템의 결과물을 수동적으로 수용하는 것을 넘어, 적극적이고 증거에 기반한 평가를 가능하게 하는 구체적인 방법론을 제공한다는 점에서 비판적 AI 리터러시의 핵심 도구로 활용할 수 있다.

AI가 제기하는 다층적이고 복합적인 문제에 대응해야 한다는 새로운 도전에 효과적으로 대응하기 위해서는 이용자들이 AI를 비판적으로 이해하고 활용할 수 있도록 지원하는 도서관과 사서의 교육적 역할 전환 또한 시급한 문제이다.

Montesi et al.(2025)의 연구에 따르면, 문헌정보학(LIS) 학생들과 현직 전문가들은 자신의 AI 리터러시 수준을 실제보다 과대평가하는 경향이 있는 것으로 나타났다. 연구진은 이러한 경향이 “인간과 유사하고 사용자 친화적인 AI 도구의 인터페이스 때문이라고 볼 수 있으며, 이는 AI 도구가 사용하기 쉽기 때문에 본인이 이 도구의 사용에 숙달했다는 잘못된 인식을 줄 수 있다”고 분석했다(Montesi et al., 2025, 11). 이는 정보전문가의 입장에서 AI의 대중화, 일상화 속에 이용자의 체계적이고 객관적인 역량 진단과 맞춤형 교육이 더욱 시급해졌음을 시사하는 것이기도 하다. 또한 현재 많은 도서관에서 제공하는 AI 관련 교육은 ChatGPT와 같은 특정 도구의 활용법이나 프롬프트 작성법에 치우치는 경향이 있다(Hervieux & Wheatley, 2024). 그러나 진정한 AI 리터러시 함양을 위해서는 알고리즘 편향, 데이터 윤리, AI 결과물에 대한 비판적 평가 방법 등 보다 근본적이고 심층적인 주제를 다루는 교육으로 전환함으로써 AI 결과물이 주는 AI의 가치관에 매몰되지 않도록 하는 노력을 지속해야 할 것이다(Hervieux & Wheatley, 2024).

AI 리터러시 교육의 성공적인 패러다임 전환은 사서 개개인의 노력만으로 이루어지기는 어려운 문제이며 개인의 노력과 함께 도서관이라는 조직 차원의 체계적인 전략과 정책이 뒷받침되어야만 가능한 과제이다.

3.4 정보전문가의 역량 및 조직의 전략

van Otterlo(2018)는 도서관은 더 이상 AI 기술을 제공하고 활용법을 안내하는 장소로만 머물러서는 안되며 이용자들이 AI 기술의 잠재적 위험을 인지하고, 이를 윤리적이고 책임감 있게 활용하도록 안내하는 ‘수호자’이자 정보 생태계의 공정성을 수호하는 ‘윤리적 게이트키퍼’로서의 역할 수행을 주장하였다(van Otterlo, 2018).

사서는 ‘윤리적 게이트키퍼’로서 단순한 정보중재를 넘어 기술에 내재된 불평등을 적극적으로 해소하고, 점점 더 자동화되는 세상에서 도서관이 지적 자유와 공정한 정보 접근을 위한 보루로 남도록 보장해야 한다. 이런 이유로 도서관은 AI 시대에 신뢰 가능한 정보 중재자로서 사회적 책무를 다하기 위한 명확한 가이드라인과 윤리적 거버넌스 체계를 구축하는 주요 역할을 해야 한다(김지연 외, 2023)고 주장하기도 한다.

이는 단순한 인적 역할 뿐 아니라 조직을 운영하는 전체적인 맥락에서 정보전문가의 역할과 조직적인 대응을 강조하는 것으로 볼 수 있다.

3.4.1 AI 시대 정보전문가의 역량

AI 시대의 정보전문가는 다면적이고 중복적인 정보기술에 대한 이해를 바탕으로 다음과

같은 역량이 필요하다는 주장이 다수로 제기되고 있다.

1) 기술적 역량과 메타 역량의 조화

Zhai(2025)는 전통적인 사서의 역량이 시스템을 운영하고 관리하는 ‘운영 능력’에 집중되었다면, 생성형 AI 시대의 사서는 AI를 효과적으로 활용하고 통제하는 ‘메타 능력(meta abilities)’을 갖추어야 한다고 주장했다(Zhai, 2025). 정확한 프롬프트 설계 능력, AI가 생성한 결과물의 사실관계 및 편향성 검증 능력, 그리고 그 과정 전반에 대한 윤리적 검토 능력 등이 포함된다.

2) 학제 간 협업 및 소통 능력

Cox(2023)는 AI 기술의 도입과 운영은 도서관 내부의 힘만으로는 불가능하기 때문에 사서는 데이터 과학자, IT 전문가, 소프트웨어 개발자, 그리고 최종 이용자 커뮤니티 사이에서 다리 역할을 하는 해석자나 번역자의 역할을 해야 한다고 주장했다(Cox, 2023). 즉, 기술 전문가에게는 도서관의 가치와 이용자의 필요를 전달하고, 이용자에게는 기술의 원리와 한계를 명확하게 설명하는 핵심적인 소통 역량을 필수적인 것으로 강조한 것이다.

3.4.2 도서관과 같은 정보조직의 전략적·정책적 방향

AI 기술을 책임감 있게 도입하고 운영하기 위한 가장 중요한 전략은 도서관/조직내에 명확한 ‘윤리적 거버넌스 체계’를 구축하는 것이다. 이 체계에 포함되어야 할 요소들은 다음과 같이 정리 가능하다.

1) 명확한 윤리 가이드라인 및 거버넌스 체계 수립

도서관은 AI 기술의 도입, 개발, 운영 전(全) 과정에 걸쳐 준수해야 할 명확한 윤리 가이드라인을 제정해야 한다. 국내외 정책 연구에서 제시된 바와 같이, 이 가이드라인에는 신뢰성, 안전성, 투명성, 위험 관리 원칙이 반드시 포함되어야 한다. AI 기술 선정 기준, 데이터 수집 및 처리 원칙, 개인정보 보호, 윤리적 고려사항 등을 명문화한 내부 규정을 마련하여 의사결정의 일관성과 투명성을 확보해야 함도 당연하다(김지연 외, 2023). 즉, 이는 AI 시대에 맞게 사서직의 ‘윤리 강령’을 재해석하고, 알고리즘의 작동을 조직의 가치 체계 안에 두는 거버넌스 체계를 구축하는 것과 같다(van Otterlo, 2018).

2) 다양한 AI 도입 전략 모색

모든 도서관이 동일한 방식으로 AI를 도입할 필요는 없다. 각 도서관은 자신의 환경, 예산, 인력, 그리고 목표에 맞는 최적의 전략을 선택해야 한다(Cox, 2022).

3) 지속적인 교육 및 재교육 프로그램 운영

사서들의 AI 역량 격차는 조직 차원의 체계적인 교육 없이는 해소되기 어렵다. 정보 전문가들이 자신의 AI 리터러시를 과대평가하는 경향이 있다는 연구(Montesi et al., 2025) 결과는 일회성 워크숍이 아닌, 지속적이고 심화된 교육 프로그램의 필요성을 시사한다. 도서관 조직은 정기적인 재교육 프로그램을 제도화하여 모든 사서가 최신 기술 동향과 윤리적 쟁점에 대한 전문성을 유지하도록 지원해야 한다.

3.4.3 알고리즘 감사 및 책임성 확보 방안

도서관이 사용하는 AI 시스템의 공정성과 책임성을 보장하기 위해서는 1회성 관리 감독이 아니라 지속적으로 유지 가능한 제도적 장치가 필요하다.

1) 공정성 감사(Fairness Audits) 도입

도서관은 도입했거나 도입할 예정인 AI 시스템의 편향성을 체계적으로 측정하고 평가하는 ‘공정성 감사’ 절차를 마련해야 할 필요가 있다(Igbinovia, 2025). 이는 알고리즘이 특정 인구 집단에게 불리하게 작용하지 않는지 정량적으로 검증하고, 감사 결과를 투명하게 공개하여 사회적 신뢰를 확보하는 최소한의 과정이라 할 것이다.

2) 책임성 있는 운영 체계 구축

AI 시스템의 결정으로 인해 이용자에게 피해가 발생했을 경우, 그 책임 소재를 명확히 하고 구제 절차를 마련하는 것이 중요하다(Narendra et al., 2025). 알고리즘의 오류나 편향으로 인한 문제 발생 시 누가, 어떻게 책임질 것인지에 대한 명확한 정책을 수립하여 공유하게 함으로써, 도서관은 기술 뒤에 숨지 않고 이용자에게 대한 책임을 다하는 신뢰받는 기관으로 남을 수 있다.

3) 인간중심의 감독 체계 유지

AI의 자동화된 결정이 이용자에게 중대한 영향을 미칠 수 있는 경우(예: 특정 정보의 필터링), 반드시 사서가 최종적으로 검토하고 개입하여 판단을 내릴 수 있는 ‘인간 중심(human-in-the-loop)’ 시스템을 보장해야 한다

(Zhai, 2025b). 이를 통해 보다 바람직한 ‘인간-AI 협업’ 문화 구축이 가능해질 것이다.

4. 결론 및 제언

본 연구는 AI 기술이 도서관 환경에 미치는 영향을 분석하고, 특히 데이터 편향성 문제에 초점을 맞추어 도서관과 사서가 직면한 윤리적 도전과 과제를 심층적으로 탐구한 연구이다.

먼저 AI 기술의 확산은 정보의 생성, 유통, 접근 방식을 근본적으로 재편하고 있으며, 도서관의 역할을 단순한 기술 수용자를 넘어 정보의 윤리적 유통을 책임지는 핵심 사회 기관으로 재정의해야 한다는 요구가 확대되고 있음을 확인했다. 본 연구는 AI가 제공하는 혁신적 기회와 동시에 ‘알고리즘 편향성’이라는 중대한 윤리적 도전을 탐구하며, 과거 분류체계의 편향성과 싸워온 도서관의 역사적 경험을 바탕으로 AI 시대의 정보 정의를 실현할 핵심 주체로서의 역할 강화를 역설하였다. AI는 참고 정보 서비스, 장서 관리 등의 자동화 및 지능화를 통해 도서관 서비스를 혁신하고 있으나, 그 이면에는 심각한 윤리적 문제가 내재되어 있음을 확인하였다. 도서관은 AI 시대의 정보 정의를 실현하는 주체로서 AI 시스템의 잠재적 편향을 체계적으로 평가하거나 데이터의 다양성을 증진할 것을 강조하였으며 이용자들이 직접 메타데이터 생성에 참여하여 기존 분류체계의 한계를 보완하는 적극적 역할 수행 지원도 주장하였다.

도서관이 이처럼 새로운 책무를 성공적으로 수행하기 위해서는 구체적인 실천 전략과 조직

적 거버넌스 구축이 필수적이다. 도서관의 교육적 역할은 단순한 AI 도구 활용법을 넘어, 이용자가 AI 생성 정보의 신뢰도를 비판적으로 평가하고 기술의 사회·윤리적 맥락을 성찰하도록 돕는 ‘비판적 AI 리터러시(Critical AI Literacy, CAIL)’ 교육의 필요성을 강조하였다.

이러한 모든 활동을 뒷받침하기 위해 조직 차원에서는 AI 도입 및 운영에 관한 명확한 윤리 가이드라인을 수립하고, 사서들을 대상으로 한 지속적인 전문가 교육을 제도화하며, AI의 자동화된 결정에 사서가 최종적으로 개입하는 ‘인간 중심(human-in-the-loop)’ 감독 체계를 포함한 책임성 있는 거버넌스를 구축해야 한다고 강조하면서 연구 분석을 마무리 하였다.

본 연구는 AI 정보기술과 리터러시에 대한 수많은 논의를 종합적으로 분석하고 그 중 중요하지만 많이 다뤄지지 않았던 ‘데이터 편향성’에 초점을 맞추어 논의를 확장시킨 연구라는 점에서 그 의의가 있다. 더불어 ‘윤리적 게이트키퍼’와 같은 개념을 사용하여 도서관을 AI 기술의 수동적 ‘소비자’가 아닌, 알고리즘 시대의 정보 정의를 실현하는 능동적 주체로 재정의했다는 점에서도 그 의의가 있다.

이러한 연구의 의의에도 불구하고 이 연구의 핵심적인 한계는 문헌 연구를 통한 내용 분석에 기반했기 때문에 국내 도서관 현장에서 AI 기술이 실제로 어떻게 도입되고 있으며, 사서와 이용자들이 어떠한 윤리적 문제에 직면하고 있는지에 대한 구체적인 현실과 요구를 사례로 충분히 반영하지 못했다는 것이다.

본 연구의 논의는 국내 AI 도서관의 도입 현황 및 사례, 거버넌스 구축에 대한 현황 분석과 같은 연구, 비판적 AI 리터러시, “비판적 **리터러시”에 대한 심층 연구, 한국형 AI 윤리 가이드라인, 거버넌스 모델 연구 등과 같은 후속 연구들을 통해 더욱 심화 발전될 수 있을 것이라 기대된다.

결론적으로, 도서관은 AI 기술의 수동적 도입을 넘어 데이터 편향성과 같은 윤리적 문제 해결을 주도하는 사회적 책무의 주체로 자리매김해야 한다. 이를 위해 다양한 주체의 비판적 **리터러시 교육을 통해 이용자와 사서의 역량을 강화하고, 투명하고 책임성 있는 윤리적 거버넌스 체계를 마련하는 것이야말로 격변하는 AI 시대에 도서관의 지속가능성을 담보하는 핵심 과제임을 강조하고자 한다.

참 고 문 헌

- 곽우정, 노영희 (2021). 도서관의 인공지능 (AI) 서비스 현황 및 서비스 제공 방안에 관한 연구. 한국도서관·정보학회지, 52(1), 155-178.
- 김지연, 석병진, 김예역, 이창훈 (2023). 안전한 AI 서비스를 위한 국내 정책 및 가이드라인 개선방안 연구. 정보보호학회논문지, 33(6), 975-988.
- 김태영, 강주연, 김건, 오효정 (2021). 지능형 기록정보서비스를 위한 선진 기술 현황 분석 및 적용

- 방안. 한국기록관리학회지, 18(4), 149-182.
- 이정미 (2023). ChatGPT, 생성형 AI 시대 도서관의 데이터 리터러시 교육에 대한 연구. 한국문헌정보학회지, 57(3), 303-323.
- 하주혜, 김성지, 오창훈 (2025). AI 리터러시가 AI 윤리에 미치는 영향: 시나리오 기반 분석. 멀티미디어학회논문지, 28(2), 333-352.
- Aithal, P. S. & Aithal, S. (2023). Effect of AI-based ChatGPT on higher education and the higher education libraries. International Journal of Management, Technology, and Social Sciences (IJMTS), 8(2), 291-312.
- Arora, A., Barrett, M., Lee, E., Oborn, E., & Prince, K. (2023). Risk and the future of AI: algorithmic bias, data colonialism, and marginalization. Information and Organization, 33(3), 100478.
- Bansal, C., Pandey, K. K., Goel, R., Sharma, A., & Jangirala, S. (2023). Artificial intelligence (AI) bias impacts: classification framework for effective mitigation. Issues in Information Systems, 24(4), 367-389.
- Berendt, B., Karadeniz, Ö., Kiyak, S., Mertens, S., & d'Haenens, L. (2023). Bias, diversity, and challenges to fairness in classification and automated text analysis. From libraries to AI and back. arXiv preprint arXiv:2303.07207.
- Blodgett, S. L., Barocas, S., Daumé III, H., & Wallach, H. (2020). Language (technology) is power: a critical survey of "bias" in NLP. arXiv preprint arXiv:2005.14050.
- Boateng, F. (2025). The transformative potential of Generative AI in academic library access services: opportunities and challenges. Information Services and Use, 45(1-2), 140-147.
- Bridges, L., Lopezosa, C., Salvadó, C., & Ferran-Ferrer, N. (2025). Moving towards critical AI literacy in LIS education: a curriculum analysis of top-ranked international programmes. The Electronic Library, 43(5), 798-817.
- BrodeFrank, J. (2025). Digital literacy & crowdsourcing: tackling descriptive and algorithmic bias through doing. Paper submitted for publication.
- Carpenter, B. (2025). The bias is inside Us: supporting AI literacy and fighting algorithmic bias. Library Trends, 73(4), 476-492.
- Cox, A. M. & Mazumdar, S. (2024). Defining artificial intelligence for librarians. Journal of Librarianship and Information Science, 56(2), 330-340.
- Cox, A. M. (2023). How artificial intelligence might change academic library work: applying the competencies literature and the theory of the professions. Journal of the Association for Information Science and Technology, 74(3), 367-380.

<https://doi.org/10.1002/asi.24635>

- Hervieux, S. & Wheatley, A. (2024). Building an AI literacy framework: perspectives from instruction librarians and current information literacy tools. *Choice*.
- Hossain, Z. (2025). School librarians developing AI literacy for an AI-driven future: leveraging the AI Citizenship Framework with scope and sequence. *Library Hi Tech News*, 42(2), 17-21.
- Igbinovia, M. O. & Mensah Danquah, M. (2025). Artificial intelligence algorithm bias in information retrieval systems and its implication for library and information science professionals: a scoping review. *Technical Services Quarterly*, 1-27.
- Islam, M. N., Ahmad, S., Aqil, M., Hu, G., Ashiq, M., Abusharhah, M. M., & Saky, S. A. T. M. (2025). Application of artificial intelligence in academic libraries: a bibliometric analysis and knowledge mapping. *Discover Artificial Intelligence*, 5(1), 59.
- James, A. B. & Filgo, E. H. (2023). Where does ChatGPT fit into the Framework for Information Literacy? The possibilities and problems of AI in library instruction. *College & Research Libraries News*, 84(9), 334.
- Li, P., Yang, J., Islam, M. A., & Ren, S. (2023). Making AI less 'thirsty': Uncovering and addressing the secret water footprint of AI models. *arXiv preprint arXiv:2304.03271*.
- Long, D. & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI conference on human factors in computing systems* (pp. 1-16). ACM.
- Montesi, M., Bautista-Puig, N., & Ferran-Ferrer, N. (2025). AILIS 1.0: A new framework to measure AI literacy in LIS. *The Journal of Academic Librarianship*, 51, 103118.
- Narendra, A. P., Dewi, C., Gunawan, L. S., & Ardi, A. S. (2025). Artificial intelligence implementation in library information systems: current trends and future studies. *Vietnam Journal of Computer Science*, 1-25.
- Ngulube, P. & Vincent Mosha, N. F. (2025). Integrating artificial intelligence-based technologies 'safely' in academic libraries: an overview through a scoping review. *Technical Services Quarterly*, 42(1), 46-67.
- Noble, S. U. (2018). *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York University Press.
- Ntoutsis, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejd, W., Vidal, M. E., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., Broelemann,

- K., Kasneci, G., Tiropas, T., & Staab, S. (2020). Bias in data-driven artificial intelligence systems—An introductory survey. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 10(3), e1356.
- Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science*, 366(6464), 447-453.
- Perrigo, B. (2023). Exclusive: The \$2 per hour workers who made ChatGPT safer. *TIME*. Available: <https://time.com/6247678/openai-chatgpt-kenya-workers/>
- Smith, C. (2022). Automating intellectual freedom: Artificial intelligence, bias, and the information landscape. *IFLA Journal*, 48(3), 422-431. <https://doi.org/10.1177/0961000620982260>
- van Otterlo, M. (2018). Gatekeeping algorithms with human ethical bias: The ethics of algorithms in archives, libraries and society. *arXiv preprint arXiv:1801.01705*.
- Varsha, P. S. (2023). How can we manage biases in artificial intelligence systems-A systematic literature review. *International Journal of Information Management Data Insights*, 3(1), 100165.
- Zhai, Y. (2025). A new paradigm of library management in the AIGC era: the construction of a three-level linkage model and its practical validation. *The Journal of Academic Librarianship*, 51(2), 103085.
- Zhang, C., Wang, B., & Ye, S. (2025). Proposing A Critical AI Literacy Framework for Academic Librarians: a case study of a database-anchored GenAI tool for Chinese studies. *International Journal of Librarianship*, 10(2), 34-47.

• 국문 참고자료의 영어 표기

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

- Ha, Ju Hye, Kim, Sung Ji, & Oh, Chang Hoon (2025). Impact of AI Literacy on AI ethics: a scenario-based analysis. *Journal of Korea Multimedia Society*, 28(2), 333-352.
- Kim, Jiyoun, Seok, Byoungjin, Kim, Yeog, & Lee, Changhoon (2023). A study on the improvement of domestic policies and guidelines for secure AI services. *Journal of The Korea Institute of Information Security & Cryptology*, 33(6), 975-988.
- Kim, Tae-Young, Gang, Ju-Yeon, Kim, Geon, & Oh, Hyo-Jung (2021). A study on the current status and application strategies for intelligent archival information services. *Journal of Korean Society of Archives and Records Management*, 18(4), 149-182.
- Kwak, Woojung & Noh, Younghee (2021). A study on the current state of the library's AI

service and the service provision plan. Journal of Korean Library and Information Science Society, 52(1), 155-178.

Lee, Jeong-Mee (2023). A study on the data literacy education in the Library of the Chat GPT, generative AI era. Journal of Korean Society for Library and Information Science, 57(3), 303-323.

