논문 2026-2-2 http://dx.doi.org/10.29056/jsf.2026.06.02

# 인공지능 기반 소프트웨어 개발 과정에서 생성된 코드의 저작권 귀속 문제

## - 실질적 유사성 판단의 새로운 과제를 중심으로 -

김시열*†

# Authorship and Copyright Ownership of Code Generated in AI-Based Software Development

Kim, Siyeol*†

요 약

인공지능 코딩 도구의 확산은 컴퓨터프로그램저작물에 대한 저작권 침해 판단, 특히 실질적 유사성 판단의 여러 과정 중 여과 단계에 새로운 부담을 더하고 있다. 인공지능이 생성한 코드는 인간의 창작성이 부재하므로 권리의 보호 대상이 되지 않기 때문에 여과 대상이 되나, 현실의 분쟁 상황에서는 코드 내용이 인간의 작성 부분과 인공지능 생성 부분이 미세하게 혼재되어 그 분리가 심히 어렵다는 문제를 지닌다. 이러한 상황을 전제로 본고에서는 인공지능 생성 코드가 여과 단계에 미치는 여러 문제들을 살펴보았고, 그 결과 문제들은 종래의 문제를 증폭하는 경우와 인공지능 활용에 따른 특유한 경우로 이원화됨을 알 수 있었다. 이의 대응을 위하여 종래 여과 절차의 인식 강화, 의거 요건의 개념 조정 및 비교 관점 확대 등의 방안을 제안하였다.

Abstract

The proliferation of AI coding tools poses a new challenge for determining copyright infringement in computer programs, particularly at the filtration step of the substantial similarity analysis. Because AI-generated code lacks human creativity, it is not protectable and is filtered out; yet in practice, human-authored and AI-generated portions are so intermingled that separating them is exceedingly difficult. The article finds the resulting problems fall into two categories: those that exacerbate existing issues and those specific to artificial intelligence. To address them, it proposes reinforcing the conventional filtration process, recalibrating the concept of reliance, and broadening the scope of comparison..

**한글키워드 :** 인공지능, 컴퓨터프로그램저작물, 여과, 창작성, 실질적 유사성

**keywords :** artificial intelligence, computer program work, filtration, creativity, substantial similarity

* 전주대학교 로컬벤처학부(법학박사)
† 교신저자: 김시열(email: sykimlaw@hanmail.net)
접수일자: 2026.06.06. 심사완료: 2026.06.15.
게재확정: 2026.06.20.

## 1. 서 론

지식재산권 분야에서 인공지능을 둘러싼 연구

가 폭넓게 이루어지고 있으나, 정작 현재의 과도기적 상황이 야기하는 실무적 문제에 대해서는 충분한 해결의 모색이 이루어지고 있다고 볼 수 없다. 소스코드를 대상으로 한 저작권 침해 판단, 특히 그중에서도 실질적 유사성 판단의 과정은 이러한 공백이 더욱 두드러지는 영역이다. 인공지능 코딩 도구[1]가 개발 실무에서 빠르게 자리잡으면서, 보호받지 못하는 표현을 여과 등의 과정을 통하여 비교 대상에서 제외하던 종래의 판단 방식에 새로운 부담이 더해지고 있기 때문이다[2].

그간 인공지능과 저작권에 관한 논의는 주로 인공지능 생성물의 저작물성과 그 권리 귀속에 집중되어 왔다[3]. 그런데 귀속의 문제는 거시적으로 '완성된 결과물의 권리자가 누구인가'를 묻는 것에 그치지 않는다. 침해 판단의 실무에서는 하나의 코드 안에서 '어떠한 표현이 권리자에게 귀속되는 보호 표현인가'를 가려내는 미시적 귀속의 문제가 더 먼저 나타나게 된다. 실질적 유사성 판단은 권리자에게 귀속되는 창작적 표현 형식만을 대비하고 그 밖의 부분을 여과하여 제외하는 작업이므로, '보호받는 표현 = 권리자에 귀속되는 표현'이라는 등식에서 보면 귀속의 문제와 여과의 문제는 동전의 양면을 이루는 것으로 이해할 수 있다[4].

이러한 점은 인공지능 생성 코드의 문제를 더욱 선명하게 드러내는데, 인공지능이 생성한 부분은 인간의 창작 요건을 결하게 되어 누구에게도 저작권이 귀속되지 않으므로, 권리자에게 귀속되는 보호 가능한 표현이 아니며 이는 여과의 대상이 된다. 그러나 현실의 코드에서는 인간이 작성한 부분과 인공지능이 생성한 부분이 미세하게 혼재되어 있어, 귀속되는 표현과 귀속되지 않는 표현을 분리해 내는 작업 자체가 매우 곤란하다. 이 논문은 바로 이러한 분리의 어려움을 침해 판단 구조의 내부에서 다루려는 것이다.

## 2. 인공지능 생성 코드와 저작권 보호범위

### 2.1 인간에 의한 표현 요건과 저작권 귀속

우리 저작권법은 컴퓨터프로그램저작물을 저작물의 한 유형으로 예시하면서, 이를 어문저작물과 같은 표현을 지닌 창작물의 일종으로 인식한다. 컴퓨터프로그램을 어문저작물과 같은 차원의 창작물로 파악하는 태도는 일찍이 CONTU 보고서에서도 확인된 바 있다. 나아가 저작권법은 저작물을 '인간의 사상 또는 감정을 표현한 창작물'로 정의하여 인간에 의하여 이루어질 것을 핵심 요건으로 삼고 있다. 그 표현을 창작한 인간인 저작자에게는 저작권이 원시적으로 귀속된다.

이러한 인간인 저작자 요건에 따르면, 인공지능이 단독으로 생성한 결과물은 그 자체로 보호대상이 되지 못한다. 따라서 누구에게도 저작권이 귀속되지 않게 된다. 인공지능을 인간이 사용하는 도구로 보아 결과물에 창작성을 인정할 수 있다는 도구론적 시각이 제시되기도 하나, 구체적 표현의 상당 부분이 인간의 의도가 아니라 모델 내부의 확률적 선택으로 결정된다는 점에서 이를 그대로 적용하기에는 한계가 있다[5]. 판례의 태도는 창작성에 관하여 완전한 의미의 독창성까지 요구하는 것은 아니지만 적어도 저작자 자신의 독자적 표현을 담고 있어야 한다고 본다[6].

### 2.2 인간–인공지능 혼합 창작에서 보호 범위의 획정 문제

현실적으로 더 문제 되는 것은 인공지능 생성물의 권리가 누구에게 귀속하는가가 아니라, 인공지능과 인간의 혼합 창작으로 이루어진 코드에서 귀속되는 표현의 범위를 어떻게 설정할 것인가이다. 실제 코드에는 인간이 직접 작성한 부분, 인공지능이 생성하여 거의 수정 없이 채택된 부

분, 인공지능의 제안을 인간이 상당히 수정한 부분, 양자가 반복적 대화를 통해 공동 형성된 부분이 혼재되어 있다. 앞의 두 유형은 귀속 여부에 큰 다툼이 없으나, 인간이 상당히 수정하거나 공동 형성한 회색지대가 문제가 된다.

저작권 등록 실무에서도 인공지능 생성 부분과 인간의 기여를 구분하여 밝히도록 하는 방향이 나타나고 있다. 결국 권리자에게 귀속되는 보호 표현과 귀속되지 않는 표현의 경계가 하나의 파일 내부에 미세하게 분포하게 되는데, 이때 실질적 유사성 판단의 전제로서 보호되는 표현을 추려내는 과정에서 인공지능이 작성한 부분을 분리해 낼 수 있는가가 관건이 된다.

### 2.3 저작권 귀속과 실질적 유사성 판단에서의 여과

저작권 침해가 인정되기 위해서는 침해자의 저작물이 권리자의 저작물에 의거하여 이를 이용하였을 것과, 두 저작물 사이에 실질적 유사성이 인정될 것이라는 두 요건이 충족되어야 한다[7]. 그리고 실질적 유사성을 판단할 때에는 창작적 표현 형식에 해당하는 부분인 권리자에게 귀속되는 보호 표현만을 가지고 대비하여야 한다. 그리고 그것이 원저작물 전체에서 차지하는 양적·질적 비중도 함께 고려하여야 한다[8]. 보호받지 못하는 표현으로 권리자에게 귀속되지 않는 표현을 비교 대상에서 제거하는 작업이 여과에 해당한다.

컴퓨터프로그램저작물을 대상으로 한 실질적 유사성 판단은 주로 추상화-여과-비교의 3단계 구조를 갖는 모델을 사용하는 경우가 많다[9]. 우리나라의 경우에는 이 모델을 공식적으로 채택한 것은 아니나 개념적으로 유사한 태도를 취하고 있는 상황으로 보인다. 이때 여과 단계에서는 아이디어에 해당하는 부분, 효율성·외부적 제약에 의해 표현이 사실상 하나(혹은 적은 수)로 수렴하여 합체(merger) 되는 부분[10], 특정 기능 구현을 위하여 불가피하게 채택되는 표현이거나 관용적인 표현, 공중의 영역에 속하는 부분 등을 제거한다. 저작권법 제101조의2가 컴퓨터프로그램 언어, 규약, 해법에는 보호가 미치지 않는다고 규정한 것도 같은 맥락이며, 이들은 모두 특정한 권리자에게 귀속될 수 없는 요소라는 점에서 공통된다.

실무에서 여과를 거친 비교 단계는 흔히 소스코드를 정량적으로 대비하여 산출한 유사도 수치에 의존한다[11]. 판례 역시 정량적 비교를 수행한 결과 서로 동일하거나 유사한 비율이 낮다는 사정 등을 들어서 실질적 유사성을 부정한 사례가 있다[12]. 앞서 언급한 바와 같이 '귀속의 획정'과 '여과'는 동전의 양면이며, 실질적으로 유사한 정도의 정량적 수치는 여과 이후 잔존하는 귀속 표현을 대상으로 산출될 때 비로소 의미를 갖는다. 인공지능의 생성 코드는 바로 이 귀속 표현의 식별을 어렵게 함으로써 여과 단계에 부담을 더하게 된다.

## 3. 인공지능 생성 코드가 여과 단계에 미치는 문제들

### 3.1 관용적 표현과 창작적 표현의 구별 곤란

표준적인 표현이나 합체, 그리고 공지 영역의 판단 기준은 표현을 선택하는 과정에서 자유로운 선택이 가능하였는가이지 인공지능이 그 표현을 얼마나 많이 사용하였는가가 아니다. 따라서 본래 창작적이던 표현이 인공지능에 의하여 빈번히 사용된다는 사정만으로 관용적 표현으로 전환되는 것으로 보는 것은 옳지 않으므로, 인공지능이 관용적 표현의 외연 자체를 확장한다고 단정하는 것은 신중하여야 한다. 문제의 본질은 관용적 표현의 범주가 넓어지는 데 있는 것이 아니라, 어떤 코드가 본래 창작성의 여지가 낮은 표준화된

표현인지 아니면 특정인에게 귀속되는 창작적 표현인지를 사후적으로 구별하기 어려워진다는 데 있다.

이 구별의 곤란성은 추상적 가능성에 그치는 것이 아니다. 한 사례를 보면, 2022년 한 개발자는 짧은 프롬프트만으로 GitHub Copilot이 자신이 권리를 갖는 코드를 출처의 표시 없이 사실상 그대로 사용하였음을 주장하였고, GitHub 측은 그것이 공개 저장소에 빈번히 등장하는 패턴에 불과하며 또한 두 코드는 서로 유사하기는 하나 전혀 다른 코드로 보아야 한다고 반박했다[13]. 즉, 동일한 표현을 두고 한쪽은 특정인에게 귀속되는 창작적 표현을 동일하게 이용한 것으로 보고, 다른 쪽은 이를 창작성의 인정 여지가 낮은 일반적으로 사용되는 표준화된 코드라고 본 것이다. 인공지능 도구 제공자 스스로도 표현물 가운데 공개 코드와 일치하거나 서로 유사한 부분이 존재한다는 점을 전제로 이를 사전에 차단하는 공개 코드 필터를 두고 있는데[14], 이는 무엇이 보호 가능한 특정한 표현이고 반대로 무엇이 표준화된 보호받을 수 없는 표현인지가 기계적으로 명확하지 않음을 역설적으로 보여준다. 결국 이 지점에서의 과제는 관용적 표현의 개념 재정의가 아니라, 그 개념을 적용하기 위한 사실인정의 곤란이라 할 수 있다.

### 3.2 여과 단위 설정 문제의 증폭

여과를 위해 코드를 분석하면 인간이 작성한 부분과 인공지능 생성 부분이 파일・함수・블록・행의 여러 수준에서 혼재되어 있어, 어느 단위에서 여과의 기준을 설정할 것인지 문제가 된다. 그런데 이 단위 설정의 어려움은 인공지능이 새롭게 창출한 것이 아니다. 여과 단위를 둘러싼 논쟁은 추상화-여과-비교 테스트가 확립된 이래 컴퓨터프로그램 저작권 분쟁에서 이미 존재하고 있던 것이며[15], 인공지능 활용의 활성화는 잠재되어 있던 이 어려움을 더욱 증폭시키는 계기가 되었다.

증폭의 이유는 인간과 인공지능의 기여가 종래보다 훨씬 미세한 수준에서 뒤섞인다는 데 있다. 한 줄의 핵심적인 수정이 함수 전체의 성격을 바꿀 수도 있고, 다수의 형식적 수정이 실질적 창작에 이르지 못할 수도 있다. 종래의 감정 실무가 함수 또는 파일 단위의 정량 비교에 익숙하다는 점을 고려하면, 인공지능이 개입한 단위와 여과의 단위가 서로 어긋날 가능성이 커지고, 그 결과 어느 표현이 권리자에게 귀속되는지를 정하는 판단의 민감도가 높아지게 된다. 여과 단위는 단순한 기술적 선택이 아니라 귀속 범위를 좌우하는 규범적 판단의 문제인 것이다.

### 3.3 학습 데이터 내재화와 의거성

의거성은 실질적 유사성과 구별되는 별개의 요건인데, 판례는 의거관계가 기존 저작물에 대한 접근 가능성과 양 저작물 사이의 유사성 등과 같은 간접증거에 의하여 추정될 수 있다고 본다[16]. 그런데 인공지능 코딩 도구는 학습 과정에서 방대한 원저작물을 내부화하기 때문에, 출력물이 학습 데이터에 포함된 특정 코드를 재현할 가능성을 갖는다. 이때의 접근은 인간이 원저작물을 보았다는 전통적인 형태가 아니라 학습이라는 형태로 추상화되어 존재한다. 더욱이 학습 데이터의 구체적 내역이 공개되지 않는 한 특정한 원저작물이 학습에 포함되었는지를 권리자가 입증하기는 매우 어렵다[17]. 학습 데이터 목록 공개를 의무화하기 위한 논의가 있으나, 아직 제도화되지 못한 상태이며 동시에 실무적으로 효용성을 얼마나 확보할 수 있을지 의문이 있다.

그런데 여기서 문제가 되는 것은 인공지능을 활용하는 소프트웨어 개발 환경으로의 변화로 인하여 여과 과정이 어려워진다는 것을 넘어, 의거성의 입증이라는 기준을 수정해야 하는 것이 아

닌가 하는 요구를 고려해야 한다는 데 있다. 접근 가능성과 유사성(현저한 유사성 등)으로부터 의거의 존재 여부를 추정하는 종래의 방식은 인간의 접근을 전제로 형성된 것인데, 인공지능이 고도로 활용된 사건에서는 그 접근이 학습으로 대체되어 기존의 판단 방식은 사실상 유의미한 작동이 불가능하게 된다. 따라서 인공지능이 생성한 코드와 관련한 사건에서는 의거성 입증을 위한 접근 요소에 원저작물의 학습 데이터 포함 가능성을 반영하고, 학습 데이터의 비공개라는 구조적 정보 비대칭을 고려하여 추정과 입증책임을 조정하는 해석이 필요한 것으로 이해된다.

### 3.4 비문언적 복제와 실질적 유사성

앞의 의거성 요건에 관한 문제와 구별되어야 할 것이 비문언적 복제의 문제이다. 이는 학습 데이터 포함 여부가 아니라, 외형이 다른 출력물이 권리자에게 귀속되는 보호받는 표현을 재현하였는지를 가리는 실질적 유사성 판단을 위한 비교의 문제이다. 본래 컴퓨터프로그램저작물에 대한 저작권 문제에서는 코드의 구조·순서·조직과 같은 비문언적 요소를 어디까지 보호할 것인가가 오래 다투어져 왔고, 추상화-여과-비교 테스트 역시 이러한 비문언적 요소를 다루기 위해 고안된 것이기도 하다. 미국의 연방대법원이 응용프로그램 인터페이스의 구조적 이용을 다룬 사건에서 보듯, 비문언적 요소의 보호 범위는 여전히 논쟁적이다[18].

인공지능 생성 코드는 변수명이나 문장 배열을 바꾸어 외형상 상이하면서도 본질적으로는 원저작물의 표현 구조를 재현할 수가 있는데, 이러한 비문언적 형태의 재현은 표현의 외형을 비교의 대상으로 하는 여과·비교 과정만으로 포착하기 어렵다. 특히 그 입증은 더욱 어렵다. 문언적인 복제의 경우에도 구체적인 입증이 요구되므로 침해의 인정이 쉽지 않은 상황에서, 표현의 외형이 변형된 비문언적 복제의 입증이 더욱 어려워질 수밖에 없다.

### 3.5 여과의 부실과 유사도 수치의 신뢰

앞서 살펴본 문제들은 결국 비교 단계에서 도출되는 정량적인 수치로 수렴한다. 적절하게 여과가 이루어지지 않게 되면, 그 왜곡이 비교 결과인 수치에 연쇄적으로 반영되는데, 그 방향은 단일하지 않게 된다. 한편으로는 권리자에게 귀속되지 않는 균질화된 표현이 유사한 것으로 계상되어 유사도 수치를 과대평가하는 경우가 있을 수 있고(거짓 양성), 다른 한편으로는 진정한 침해 부분이 관용적인 표현 등으로 오인되어 제외됨으로써 유사도 수치를 과소하게 평가하는 문제가 있을 수 있다(거짓 음성).

정량적 유사도 수치가 실무적으로 실질적 유사성 판단의 중요한 지표로 기능하는 현실에서, 여과가 형식적으로 이루어진다면 감정 결과의 신뢰성은 근본적으로 흔들리게 된다. 유사도의 과대평가는 정당한 개발 활동을 침해로 보게 되는 위험을 지니고, 과소평가는 실제 침해에 해당하는 행위에 면책을 주는 위험을 각각 키우게 된다. 두 위험 모두 법적 안정성과 예측 가능성을 훼손하게 된다. 따라서 인공지능 환경에서는 정량적인 수치 그 자체보다 그 수치가 어떠한 여과를 거쳐 산출되었는지가 더욱 중요한 검토 대상이 된다.

## 4. 문제의 개선방안

### 4.1 여과 절차의 위상 재정립

판례가 이미 창작적 표현만을 대비의 대상으로 하고 있다는 점에서 보호의 대상이 되는 표현만을 비교하여 실질적 유사성을 판단한다는 원칙은 이미 확립되어 있다. 그러나 인간과 인공지능의 기여가 미세하게 혼재된 코드에서는 ‘무엇이

권리자에 귀속되는 보호 가능한 표현인가'를 가리는 작업이 별도의 판단으로 의식되지 못한 채 여과 과정에서 형식적으로 처리될 위험이 높다. 따라서 이 기존의 원칙을 인공지능 활용으로 변화된 환경에 맞추어 조정하는 것이 필요하다.

구체적으로는 우리 판례의 태도에서 명확히 채택한 것은 아니지만 추상화-여과-비교의 3단계 테스트 단계 중, 추상화를 통해 분석 층위를 정한 뒤 곧바로 여과로 나아가기 전에, 각 표현들이 인간의 창작인지, 인공지능의 생성물인지, 혹은 양자의 혼합물인지를 먼저 판단하고 그 귀속 주체를 확정하는 단계를 마련하는 것이 필요하다고 생각한다. 인공지능이 단독으로 생성한 부분은 누구에게도 귀속되지 않으므로 이 단계에서 보호 대상 제외가 이루어지고, 이어지는 과정인 여과에서는 나머지 부분에 대하여 종래의 기준을 적용하여 진행할 수 있다.

### 4.2 의거 요건의 개념적 조정

의거는 실질적 유사성과 구별되는 별개의 요건으로, 판례는 원저작물에 대한 접근 가능성과 양 비교 대상 사이의 유사성(현저한 유사성 등) 등의 간접증거를 통하여 의거 관계의 인정 여부를 추정한다.

인공지능 활용이 개입된 사건에서는 이 접근이 인간이 원저작물을 보았다는 전통적인 형태가 아니라 학습이라는 형태로 추상화되어 존재하기 때문에, 접근 가능성을 매개로 한 종래의 추정 도식이 그대로 작동하지 않는다. 그렇다고 하여 학습 데이터에 포함되었을 가능성만으로 의거를 곧바로 인정할 수도 없는 것이다. 대규모 언어 모델은 사실상 공개된 코드 전체를 학습하므로, 단순히 포함만으로 의거를 추정하게 되면 의거성은 거의 모든 사건에서 자동적으로 인정되어버리는 현상이 발생할 수도 있게 된다. 이렇게 되면 의거성은 요건으로서의 기능을 상실하게 된다. 따라서 학습 데이터가 포함되었다는 사정은 의거성이 있는지를 판단하는 것의 출발점일 뿐, 그 자체로 종착점이 될 수는 없다.

이 문제는 의거의 추정을 하나의 사실로 단정하지 않고 단계적으로 구성함으로써 완화할 수 있다고 본다. 예를 들어, 원저작물이 학습 데이터에 포함되었을 상당한 개연성이 인정되는 단계라면 의거의 약한 추정이 가능하며, 이에 더하여 양자 사이에 우연으로는 설명하기 어려운 현저한 유사성 등이 인정된다면 의거의 강한 추정으로 구성할 수 있다. 한편, 이용자의 반대 입증을 통해서 이 추정은 번복될 수 있다. 요컨대 포함과 재현의 결합 정도에 따라 추정의 강도를 달리하고 그 번복 사유가 명시되는 구조이다.

이러한 단계 구조의 작동을 위해서는 입증 환경의 조정이 함께 이루어져야 한다. 학습 데이터의 내역이 공개되지 않는 한 첫 단계의 개연성조차 권리자가 입증하기 어렵기 때문이다. 따라서 이와 같은 문제 해결을 위한 제도적 개선이 함께 이루어져야 한다.

### 4.3 비교의 관점을 확대

전통적인 비교 방식은 소스코드와 같은 어문저작물의 성격을 지닌 표현물을 주된 비교 대상(보호 대상)으로 하고 있다. 그러나 인공지능의 활용으로 인한 환경 변화는 이와 같은 협소한 저작권 보호 범위로는 보호의 실효성을 확보할 수 없을 뿐 아니라, 보호의 효율 등을 저해할 수 있다. 따라서 코드의 구조·순서·조직과 같은 비문언적 표현의 재현까지도 실질적 유사성 판단을 위한 비교 대상이 될 수 있도록 비교의 층위를 넓혀야 할 필요가 있다.

대규모 언어 모델은 학습한 표현을 표면적 토큰 단위로 복제하기보다 그 구조적 표상을 바탕으로 재생성하므로, 변수명과 문장 배열 등 외형을 달리하면서도 구조·조직을 보존하는 환언적

재현이 기본적으로 발생하게 된다. 인간에 의한 비문언적 모방이 상당한 의도와 노력을 필요로 하였던 것과 달리, 인공지능에 있어서는 이러한 비문언적 재현이 손쉽고 빈번하게 발생한다는 점에서 문제의 양상이 달라지는 것이다. 따라서 비문언적 표현에 대한 범위까지 보호의 범위를 확장하고 이들까지 비교의 대상으로 넓히게 됨으로써 보호가 어려운 환경에서 보호의 강도를 높일 수 있는 수단이 될 수 있다.

다만, 비교 관점을 이렇게 확장하는 것이 보호 범위 확장으로 바로 이어져서는 곤란하다. 비문언적 요소에는 아이디어·기능·효율성 등에 의하여 지배되는 영역이 혼재되어 있기 때문에 비문언적 표현에 있어서 저작권 보호 범위의 논란은 여전히 해결되지 않고 있기 때문이다. 이에 비교 대상의 확대와 보호 대상의 확대를 연결하는 법 논리의 고민이 필요하다.

## 5. 결 론

지금까지 인공지능 생성 코드의 저작권 귀속 문제를 일반적으로 다루어왔던 권리의 귀속이 아니라 실질적 유사성 판단에서 권리자에게 귀속되는 보호되는 표현을 가려내는 여과 문제의 시각에서 살펴보았다. 그리고 보호받는 표현과 그렇지 않은 표현 간의 구별이 곤란해지고 있으며 여과 단위 설정과 같은 종래의 문제들이 인공지능의 활용으로 인하여 증폭되고 있으며, 학습 데이터 내재화에 따른 의거성 교란과 비문언적 표현의 실질적 유사성 판단 곤란(식별 곤란)은 인공지능 활용으로 인하여 발생하는 특유한 문제로 볼 수 있었다.

인공지능이 코드의 개발 과정에 굉장히 적극적으로 관여하기 시작하면서 컴퓨터프로그램을 만들어내는 환경에는 큰 변화가 발생했다. 특히 인간의 창작물만을 보호하는 저작권법 체계에서 컴퓨터프로그램저작물의 보호 문제는 굉장히 혼란스러운 논의로 이어질 수밖에 없다. 이의 해결을 위해 본고에서는 몇 가지 대안을 제시하였는데, 이들 대안들이 그 자체로 의미를 지닐 수는 없으며, 이를 조력하는 정보제공 확보 등의 제도적 개선이 함께 이루어져야 의미를 가질 수 있을 것이다.

인공지능 시대의 과제는 새로운 침해 판단 체계를 창설하는 것이 아니라, 기존의 법리가 딛고 선 전제를 재구성하는 데 있다고 본다. 이를 어떻게 해결해 갈 것인가는 앞으로의 과제로 남는다.

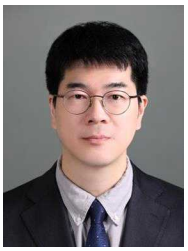
이 연구는 2025학년도 전주대학교
학술연구지원사업의 지원을 받아 수행되었음

## 참 고 문 헌

[1] GitHub, “GitHub Copilot · Your AI pair programmer”, https://github.com/features/copilot
[2] Kim Si Yeol, "Study on Quantitative Analysis for Substantial Similarity Determination of Computer Programs", The Journal of Intellectual Property, 6(4), pp.65-96, 2011.12,
[3] Pamela Samuelson, “Generative AI Meets Copyright”, 381 Science, pp.158-161, 2023; Jane C. Ginsburg & Luke Ali Budiardjo, “Authors and Machines”, 34 Berkeley Tech. L.J. pp.343-448, 2019.
[4] Mark A. Lemley & Bryan Casey, “Fair Learning”, 99 Texas Law Review, pp.743-785, 2021
[5] Jane C. Ginsburg & Luke Ali Budiardjo, “Authors and Machines”, 34 Berkeley Tech. L.J. pp.343-448, 2019.

[6] Supreme Court Judgment 2017Da212095 Sentenced June 27, 2019
[7] Supreme Court Judgment 2019Da268061 Sentenced June 30, 2021
[8] Supreme Court Judgment 2019Da268061 Sentenced June 30, 2021
[9] Computer Associates Int'l, Inc. v. Altai, Inc., 982 F.2d 693 (2d Cir. 1992).
[10] Peter S. Menell, "An Analysis of the Scope of Copyright Protection for Application Programs", 41 Stanford Law Review, p.1045, 1989
[11] Kim Si Yeol, "Study on Quantitative Analysis for Substantial Similarity Determination of Computer Programs", The Journal of Intellectual Property, 6(4), pp.65-96, 2011.12,
[12] Supreme Court Judgment 2011Do626 Sentenced January 27, 2012
[13] "GitHub Copilot may steer Microsoft into a copyright lawsuit", The Register, 2022. 10. 19.
[14] GitHub, "GitHub Copilot · Your AI pair programmer", https://github.com/features/copilot
[15] Computer Associates Int'l, Inc. v. Altai, Inc.,
[16] Supreme Court Judgment 2005Da35707 Sentenced December 13, 2007
[17] B.L.W. Sobel, "Artificial Intelligence's Fair Use Crisis", 41 Columbia Journal of Law & the Arts, pp.45-97, 2017.
[18] Google LLC v. Oracle America, Inc., 141 S. Ct. 1183 (2021).

## 저 자 소 개



김시열(Kim, Siyeol)

2012.8 숭실대학교 대학원, 법학박사
2007.6-2012.6 한국저작권위원회
2012.6-2024.2 한국지식재산연구원 연구위원
2024.3-현재 전주대학교 로컬벤처학부, 부교수
<주관심분야> 실질적 유사성, 소프트웨어 계약 분쟁 등