

# 저작권 이슈 트렌드



COPYRIGHT ISSUE TREND



한국저작권위원회  
KOREA COPYRIGHT COMMISSION

# CONTENTS

## 저작권 이슈 트렌드

Biweekly Report | 통권 제80호(2026. 4-2)

- 위키피디아의 AI 금지 정책과 AI 작성 텍스트 탐지 기술
- 영국 PRS, 스티브 운영사 상대로 음악 저작권 침해 소송 제기
- 마이크로소프트, 음성·이미지 생성에 특화된 자체 AI 모델  
3종 공개



# 위키피디아의 AI 금지 정책과 AI 작성 텍스트 탐지 기술

## 뉴스 브리프

2026년 3월, 온라인 백과사전 위키피디아는 LLM을 활용한 기사 작성 및 재작성을 전면 금지하는 정책을 도입했다. 다만, 문장 교정과 번역은 제한적으로 허용될 예정이다. 편집자 투표에서 40:2의 압도적 찬성으로 통과된 이 정책은 LLM이 위키피디아의 핵심 콘텐츠 정책을 위반한다는 판단에 근거한다. 위키미디어 재단은 AI의 확산으로 사이트 방문자 수가 전년 대비 8% 감소했다고 밝혔다. 일부 AI 에이전트는 편집 제안이 거절된 후 개발자에 대한 비판적 텍스트를 생성하거나, 차단 조치 후 스스로 활동을 일시 중단하는 등 예측 불가능한 작동 양상을 보였다. 이러한 상황은 AI 산출 텍스트와 인간 창작물을 구별하는 기술적 검증 수단의 필요성을 부각시킨다. 본 보고서는 기계학습 분류 모델의 텍스트 탐지 메커니즘을 분석하고, 저작권 보호 맥락에서의 적용 가능성을 검토한다.

## 뉴스 플러스

### I. 서론 : 위키피디아 AI 금지 정책과 텍스트 검증 기술의 부상

#### • 위키피디아의 AI 생성 콘텐츠 전면 금지 정책 배경

2026년 3월, 세계 최대 온라인 백과사전 사이트 위키피디아(Wikipedia)는 대규모 언어모델(Large Language Model, 이하 LLM)을 이용한 기사 작성 및 재작성을 전면 금지하는 정책을 공식 도입했다. 이 정책은 위키피디아 편집자 커뮤니티의 투표를 거쳐 40:2라는 압도적 찬성으로 통과되었으며, 챗GPT(ChatGPT), 제미니(Gemini), 딥시크(DeepSeek) 등 주요 LLM이 위키피디아의 핵심 콘텐츠 정책인 검증 가능성, 중립적 관점, 독창적 연구 금지 원칙을 위반한다는 판단에 기초한다. 한편, 위키피디아의 소유주인 위키미디어 재단(Wikimedia Foundation)은 생성형 AI의 확산이 사이트 신뢰도에 영향을 미쳤으며, 실제로 사이트 방문자 수가 전년 대비 8% 감소했다고 밝혔다. 이번 정책은 단순히 기술 사용을 제한하는 차원을 넘어, 온라인 지식 생산 과정에서 인간의 책임을 재확인하고 확대하려는 시도로 해석된다.



위키피디아의 AI 금지 정책 도입 배경에는 AI 에이전트의 누적된 문제 행동 사례가 있다. 위키피디아는 위키프로젝트 AI 클린업(WikiProject AI Cleanup)이라는 프로젝트를 통해 AI가 생성한 저품질 콘텐츠를 식별하고 제거하는 작업을 진행해 왔으며, 이 과정에서 AI 봇이 출처 목록을 허구로 작성하거나 다른 출처를 표절하는 사례를 다수 발견했다. 특히 톰-어시스턴트(Tom-Assistant)라는 AI 에이전트는 승인 절차를 건너뛰고 스스로 판단한 주제를 선정해 기사를 작성했으며, 위키피디아 페이지 편집 권한이 차단된 이후에는 개발자에 대한 비판적인 내용의 블로그 글을 게시하는 등 예측 불가능한 작동 양상을 보였다. 이러한 사례는 AI가 단순히 기술적 오류를 넘어, 온라인 커뮤니티의 규범과 충돌하는 방식으로 작동할 수 있음을 보여준다.

## II. 본론: 해석 가능한 AI 텍스트 분류 모델의 작동 원리

### • 기계학습 분류기의 텍스트 판별 메커니즘

AI 산출 텍스트를 탐지하는 기술은 기본적으로 기계학습 분류기를 활용한다. 분류기는 주어진 텍스트를 인간 작성과 AI 산출 두 가지로만 구분하는 이진 분류 작업을 수행하며, 다양한 기계학습 알고리즘이 적용된다. 이들 알고리즘은 텍스트를 수치화된 특징 벡터로 변환한 뒤, 고차원 특징 공간에서 인간 작성 텍스트와 AI 산출 텍스트를 구분하는 경계면을 학습한다. 주목할 점은 단순한 선형 모델조차 매우 높은 정확도를 달성한다는 사실이다. 이는 알고리즘의 복잡성보다 어떤 특징을 추출하느냐가 분류 성능에 더 큰 영향을 미친다는 것을 의미한다. 훈련 과정에서는 홀드아웃 검증\* 방식을 적용하여, 훈련 데이터로 모델을 학습시키고 검증 데이터로 성능을 평가하여 최적의 분류기를 선택한다.

\* 홀드아웃 검증(holdout validation): 전체 데이터를 훈련용과 검증용으로 나누어, 훈련 데이터로 모델을 학습시키고 검증 데이터로 성능을 평가하여 최적의 분류기를 선택하는 방법

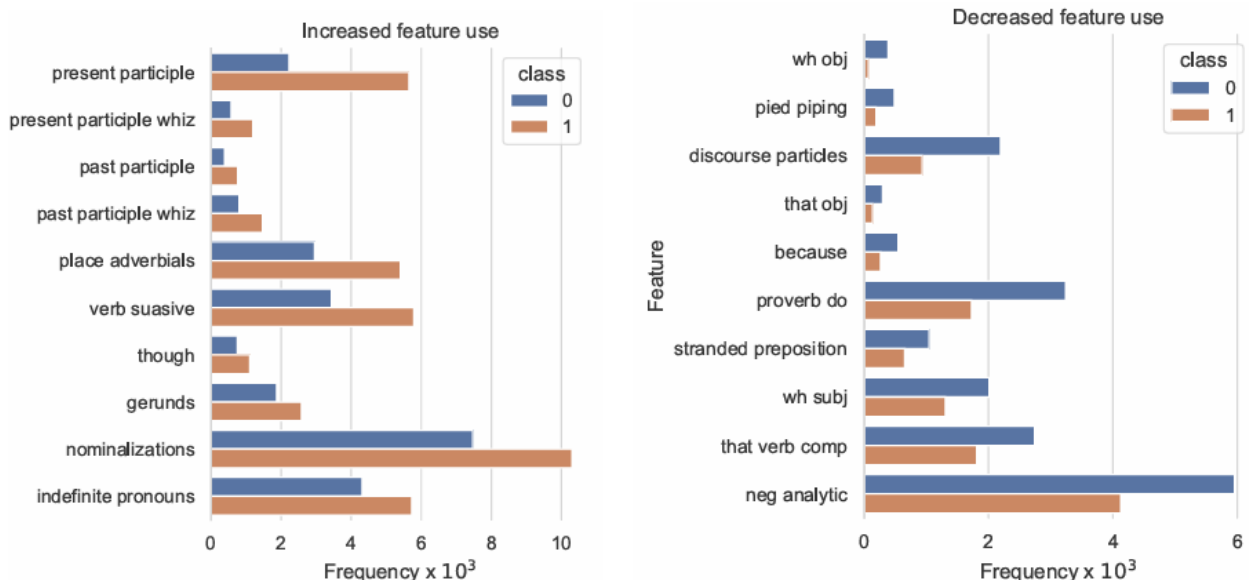
그러나 분류 정확도만으로는 실용성을 담보하기 어렵다. 전통적인 접근법은 단어 빈도, 문장 길이, 구두점 사용 패턴 등 표면적 통계 정보를 특징으로 활용했으나, 이러한 특징만으로는 LLM이 생성한 정교한 텍스트를 구별하기 어렵다. 최근에는 딥러닝 기반 임베딩 기법을 통해 의미론적 특징을 추출하거나, 문맥 속의 단어 관계를 분석하는 방식으로 성능을 개선하고 있다. 그러나 이러한 특징은 사람이 직관적으로 이해하기 어려워, 모델이 왜 특정 텍스트를 AI 산출물로 판단했는지 설명하기 곤란하다는 문제가 있다. 이는 법적 분쟁이나 저작권 검증 과정에서 치명적인 약점으로 작용한다. 판단 과정이나 근거를 알 수 없고, 단순히 확률값만 제시하는 블랙박스 모델로는 법정에서 저작권 침해나 AI 사용의 증거로 채택되기 어렵고, 편집자나 심사자가 판단 근거를 검토할 수 없기 때문이다. 따라서 모델의 판단 근거를 투명하게 제시할 수 있는 해석 가능성이 핵심 요건으로 부상하고 있으며, 이를 위해서는 인간이 직접 확인하고 평가할 수 있는 언어적 특징 추출이 필수적이다.

• LLM 산출 텍스트의 언어적 지문과 특징 추출

LLM은 텍스트를 생성할 때 특유의 언어적 지문을 남긴다. 가장 두드러진 패턴은 재표현 경향이다. LLM은 학습 데이터에서 다양한 동의어와 표현을 학습하며, 텍스트 생성 시 빈번하게 사용되는 단어를 동의어로 대체하는 경향을 보인다. 예를 들어 평범한 동사는 보다 격식 있거나 문학적인 동사로, 일상적인 형용사는 더 현학적이거나 고급스러운 형용사로 치환된다. 이는 LLM이 학습 과정에서 문학 작품이나 전문 문서를 다량으로 학습했기 때문에, 일상 대화체보다 격식체를 선호하는 경향으로 나타난다. 또한 단순한 문장 구조를 복잡한 구문으로 바꾸거나, 일상적 명사를 전문 용어로 대체하는 등 전반적으로 문장을 고급화하려는 패턴이 관찰된다.

이러한 패턴은 체계적으로 범주화될 수 있다. 연구자들은 대화 묘사 동사, 일반 서술 동사, 전치사, 부사, 형용사, 문법 구조, 명사 등 7개 하위 범주로 분류하여 약 200개 사례를 식별했다. 각 범주는 LLM의 서로 다른 언어 처리 특성을 반영한다. 대화 묘사 동사 범주에서는 발화 행위를 표현하는 방식의 차이가 드러나며, 일반 서술 동사 범주에서는 사고나 행동을 묘사하는 어휘 선택의 차이가 포착된다. 문법 구조 범주에서는 단순 시제가 완료 시제로, 직접 표현이 간접 표현으로 전환되는 등 구문 복잡도 증가 패턴이 나타난다. 이 외에도 시대적 표기 변화를 반영하지 못하는 패턴이 중요한 단서가 된다. LLM이 20세기 초반 이전의 문헌을 학습했을 경우, 당시 사용되던 맞춤법이나 띄어쓰기 방식을 현대 텍스트에 부적절하게 삽입하는 오류를 범한다. 또한 영국식 철자와 미국식 철자를 섞어서 쓰거나, 특정 지역 방언 표현을 맥락 없이 사용하는 등의 문제도 발견된다. 다음은 인간 작성 텍스트와 LLM 산출 텍스트의 사용 어휘 및 표현 빈도 차이를 나타낸 그래프다.

[그림] 인간 작성 텍스트(파랑)와 LLM 산출 텍스트(주황)의 사용 어휘 및 표현 빈도 차이



출처: Minerva Suvanto 외 3명, "Interpretable Text Classification Applied to the Detection of LLM-generated Creative Writing", arXiv, 2026.01.12., <https://arxiv.org/pdf/2601.07368>



특징 추출 과정은 대규모 텍스트 데이터 분석과 전문가 판단을 결합하여 이루어진다. 연구자들은 대규모 텍스트 데이터베이스에서 특정 단어나 구문의 사용 빈도 변화를 추적하고, 이를 LLM 학습 데이터의 시간적 분포와 비교한다. 구글 엔그램 뷰어(Google n-gram Viewer)\*와 같은 도구를 활용하면 특정 표현이 역사적으로 어떻게 사용되었는지 시각화할 수 있으며, 이를 통해 LLM이 구식 표현을 현대적 맥락에서 부적절하게 사용하는 패턴을 포착할 수 있다. 동시에 언어학 전문가가 개별 사례를 검토하여 해당 특징이 실제로 AI 산출 텍스트를 구별하는 유의미한 단서인지 검증한다. 이렇게 확보된 특징은 분류 모델의 입력값으로 사용되며, 모델의 판단 근거를 사람이 이해할 수 있는 형태로 제시할 수 있게 한다. 핵심은 통계적 수치가 아니라 구체적인 언어 요소를 제시함으로써, 편집자나 법률 전문가가 직접 해당 특징을 확인하고 평가할 수 있도록 한다는 점에 있다.

\* 구글 엔그램 뷰어(Google n-gram Viewer): 구글에서 개발한 검색 엔진으로, 1500년부터 2019년까지 출판된 인쇄물에서 발견된 n-gram의 연간 수를 사용하여 검색 문자열 집합의 빈도를 차트로 나타냄. n-gram은 연속된 텍스트 데이터에서 인접한 n개의 요소를 하나의 토큰으로 추출하여 분석하는 통계적 언어 모델링 기법임

#### • 저작권 보호 맥락에서의 적용 가능성과 한계

해석 가능한 AI 텍스트 탐지 기술은 저작권 보호 맥락에서 두 가지 방향으로 활용될 수 있다. 첫째, 창작물의 작성 주체를 검증하는 도구로 기능한다. 저작권 분쟁에서 특정 작품이 인간의 창작물인지 AI 산출물인지 판단해야 하는 경우, 탐지 모델은 언어적 특징 분석을 통해 객관적 근거를 제시할 수 있다. 예를 들어 소설 원고에서 대화 묘사 동사의 재표현 패턴이 집중적으로 발견되거나, 시대적으로 부적절한 구식 표기가 반복된다면 이는 LLM 개입 가능성을 시사하는 구체적 단서가 된다. 둘째, 온라인 플랫폼의 콘텐츠 검증 시스템으로 활용될 수 있다. 위키피디아와 같은 협업 기반 플랫폼에서는 대량으로 유입되는 기여 내용을 실시간으로 검토해야 하며, 자동화된 탐지 모델이 의심스러운 텍스트를 1차 필터링하고 인간 편집자가 최종 판단하는 방식으로 운영될 수 있다. 이는 AI 산출 콘텐츠의 무분별한 확산을 억제하고, 인간 창작자의 기여를 보호하는 기술적 보호 장치로 작동한다.

그러나 이러한 기술은 한계와 위험성을 동시에 안고 있다. 가장 큰 문제는 오탐지 가능성이다. LLM이 선호하는 언어적 특징이 인간 작성자의 문체와 우연히 일치할 경우, 정당한 창작물이 AI 산출물로 오인될 수 있다. 특히 격식체를 선호하거나 문학적 표현을 즐겨 사용하는 작가의 경우 이러한 가능성이 높아진다. 또한 LLM 기술이 빠르게 발전하면서 탐지 모델이 식별하는 언어적 지문 자체가 시간이 지나면 무용해질 가능성도 있다. 차세대 LLM이 재표현 패턴을 최소화하거나, 시대적 표기 오류를 자동으로 보정하는 기능을 갖추게 되면 해당 탐지 기법의 효과는 감소한다. 이 외에도 표현의 자유 침해 우려가 제기된다. 탐지 모델이 특정 문체나 어휘 선택을 AI 산출의 증거로 간주할 경우, 작가들이 자신의 표현 방식을 의도적으로 제한하는 위축 효과가 발생할 수 있다. 따라서 기술적 검증은 최종 판단 도구가 아니라 보조 수단으로 활용되어야 하며, 인간 전문가의 종합적 판단과 커뮤니티 기반 검증 절차가 병행되어야 한다.

## 결론 : 기술적 검증과 커뮤니티 거버넌스의 공존

### • AI 시대 콘텐츠 신뢰성 확보를 위한 과제

위키피디아의 AI 작성 금지 정책은 온라인 지식 생태계가 직면한 근본적 딜레마를 드러낸다. LLM 기술이 고품질 텍스트를 빠르게 생성할 수 있게 되면서, 콘텐츠 생산의 효율성은 극적으로 향상되었지만 동시에 진위 판별과 책임 소재 확인이라는 새로운 과제가 등장했다. 해석 가능한 AI 텍스트 탐지 기술은 이러한 과제에 대한 기술적 해법을 제시한다. 언어적 특징 분석을 통해 AI 산출 텍스트를 식별하고, 그 판단 근거를 인간이 이해할 수 있는 형태로 제시함으로써 저작권 검증과 콘텐츠 품질 관리에 실질적으로 기여할 수 있다. 그러나 이는 완전한 해결책이 아니라 보조 수단에 불과하다. 기술적 검증만으로는 오탐지 위험과 표현의 자유 침해 우려를 완전히 해소할 수 없으며, 결국 인간 전문가의 판단과 커뮤니티 기반 검증 절차가 병행되어야 한다. 위키피디아 사례가 보여주듯, 기술적 도구와 인간 편집자의 협력, 그리고 투명한 정책 수립 과정이 결합될 때 비로소 신뢰할 수 있는 콘텐츠 생태계가 유지될 수 있다.

향후 과제는 크게 세 가지로 전망된다. 첫째, 탐지 기술의 지속적인 고도화가 필요하다. LLM 기술이 빠르게 진화하면서 기존 언어적 지문이 사라지거나 변형될 가능성이 높으므로, 탐지 모델 역시 새로운 패턴을 지속적으로 학습하고 업데이트해야 한다. 둘째, 제도적 보완이 뒷받침이다. AI 산출 텍스트의 표시 의무화, 플랫폼의 검증 책임 강화, 오탐지 피해자 구제 절차 마련 등 법적·정책적 틀이 함께 정비되어야 기술적 검증이 실효성을 가질 수 있다. 셋째, 산업적 대응 방향의 재정립이 요구된다. 창작 산업에서는 AI 도구를 단순히 배제하는 대신, 인간 창작자와 AI의 협업 방식을 명확히 정의하고 기여도를 투명하게 표시하는 방향으로 나아가야 한다. 이를 통해 기술 혁신의 이점을 활용하면서도 인간 창작자의 권리를 보호하고, 콘텐츠 신뢰성을 유지하는 균형점을 찾을 수 있을 것이다.

### 참고문헌

- Lucas Ropek, "Wikipedia cracks down on the use of AI in article writing", TechCrunch, 2026.03.26., <https://techcrunch.com/2026/03/26/wikipedia-cracks-down-on-the-use-of-ai-in-article-writing/>
- Danny Bradbury, "Wikipedia's AI agent row likely just the beginning of the bot-ocalypse", Malwarebytes, 2026.04.01., <https://www.malwarebytes.com/blog/ai/2026/04/wikipedias-ai-agent-row-likely-just-the-beginning-of-the-bot-ocalypse>
- Cathy Resurreccion, "AI writing faces restrictions under new Wikipedia rules", CoinGeek, 2026.04.06., <https://coingeek.com/ai-writing-faces-restrictions-under-new-wikipedia-rules/>
- Minerva Suvanto 외 3명, "Interpretable Text Classification Applied to the Detection of LLM-generated Creative Writing", arXiv, 2026.01.12., <https://arxiv.org/pdf/2601.07368>



## 기술용어

순번	용어	설명
1	홀드아웃 검증 (holdout validation)	전체 데이터를 훈련용과 검증용으로 나누어, 훈련 데이터로 모델을 학습시키고 검증 데이터로 성능을 평가하여 최적의 분류기를 선택하는 방법
2	구글 엔그램 뷰어 (Google n-gram Viewer)	구글에서 개발한 검색 엔진으로, 1500년부터 2019년까지 출판된 인쇄물에서 발견된 n-gram의 연간 수를 사용하여 검색 문자열 집합의 빈도를 차트로 나타냄. n-gram은 연속된 텍스트 데이터에서 인접한 n개의 요소를 하나의 토큰으로 추출하여 분석하는 통계적 언어 모델링 기법임



# 영국 PRS, 스팀 운영사 상대로 음악 저작권 침해 소송 제기

## 뉴스 브리프

영국의 저작권 관리 단체 PRS(Performing Right Society)는 2026년 3월 세계 최대 PC 게임 유통 플랫폼 스팀(Steam)의 운영사 밸브(Valve)를 상대로 음악 저작권 침해 소송을 제기하였다. PRS는 밸브가 2003년 서비스 출시 이후 약 23년간 게임 내 음악의 이용제공(making available)에 관한 라이선스를 취득하지 않았다고 주장하였으며, 영국 저작권 체계에서는 게임 개발사가 취득하는 싱크 라이선스와는 별도로 유통 플랫폼 역시 이용제공 행위에 대한 라이선스를 취득해야 한다는 점을 강조하였다. 이번 PRS의 밸브 소송 사례는 디지털 콘텐츠 유통 구조가 고도화됨에 따라 저작권의 권리 유형별 책임 부담 주체 역시 세분화되고 있음을 시사한다.

## 뉴스 플러스

### 1. 서론 : 게임 유통 플랫폼과 음악 저작권의 충돌

#### • PRS v. 밸브 소송 사건 개요

2026년 3월 4일, 영국의 저작권 관리 단체 PRS(Performing Right Society)는 세계 최대 PC 게임 유통 플랫폼 스팀(Steam)의 운영사인 밸브 코퍼레이션(Valve Corporation, 이하 밸브)에 대한 소송을 제기하였다. PRS는 밸브가 2003년 스팀 출시 이후 유통해 온 게임 내 음악 저작물에 대한 적절한 라이선스를 취득하지 않은 채 운영을 지속해 왔다고 주장하였다.

PRS는 수년간 밸브 측에 적절한 라이선스를 취득할 것을 요청하였으나, 실질적인 라이선스 취득 절차는 진행되지 않은 것으로 확인된다. 이에 PRS는 영국 1988 저작권, 디자인 및 특허법(Copyright, Designs, and Patents Act 1988, 이하 CDPA) 제20조<sup>1)</sup>에 의거하여 소송을 제기하였다. PRS는 밸브가 과거

라이선스 미취득분에 대해서는 소급하여 재정리하고, 향후 이용분에 대해서도 별도의 라이선스를 체결할 것을 요청하였다.

## II. 본론: PRS v. 밸브 소송 사건의 쟁점

### • 영국 게임 음악 저작권 체계의 이해

이번 PRS v. 밸브 소송 사건을 이해하기 위해서는 영국 음악 저작권 체계의 권리 구조를 파악할 필요가 있다. 통상 게임 개발사나 퍼블리셔는 자사 게임에 음악을 삽입하기 위해 음악 출판사로부터 싱크 라이선스(sync license)\*를 취득한다. 이는 TV나 영화에서의 음악 사용과 유사한 방식이다.

\* 싱크 라이선스(sync license): 저작권이 있는 음악을 영상물이나 게임 등 시청각물과 동기화하여 삽입·이용할 수 있도록 허락하는 권리처리 방식

그러나, 영국 CDPA상 싱크 라이선스의 핵심인 ‘음악을 게임에 삽입하는 행위’는 복제권(제17조2)과 관련된 문제로 이해되며, 이후 온라인 플랫폼을 통한 유통은 별도의 공중전달권·이용제공권(제20조) 관련 문제로 구별될 수 있다. 즉, 게임 개발사나 퍼블리셔가 싱크 라이선스를 취득하는 것과 별개로, 유통 플랫폼인 스팀이 이용제공권 관련 라이선스를 취득해야 한다는 것이다. 한편, 영국에서 이용제공에 관한 권리는 음악 출판사가 아닌 PRS만이 라이선스할 수 있는 영역으로 분류된다.

[표] 영국 게임 음악 환경에서 싱크 라이선스와 이용제공권 비교

구분	싱크 라이선스 (sync license)	이용제공권 (making available right)
대상 행위	음악을 게임에 삽입하는 행위	음악이 포함된 게임을 다운로드·스트리밍으로 이용자에게 제공하는 행위
주요 관련 권리 유형	복제권	공중전달권
권리 보유 주체	음악 출판사	PRS (작곡가·작사가가 가입 시 PRS에 양도)
라이선스 취득 주체	게임 개발사/퍼블리셔	게임 유통 플랫폼 운영사
라이선스 체결 상대	개별 음악 출판사와 직접 협상	PRS를 통해 포괄적으로 체결
적용 시점	게임 제작 단계	게임 유통 단계

출처: 참고문헌을 종합하여 재구성

### • 게임 유통 플랫폼으로서 스팀의 입지

스팀은 PC 게임 유통 시장에서 약 75%의 점유율을 차지하고 있으며, 월간 활성 사용자 수는 1억 4,700만 명에 달한다<sup>3)</sup>. 영국의 PC 게임 전문 매체 피시 게이머(PC Gamer)는 스팀을 “이용자들이

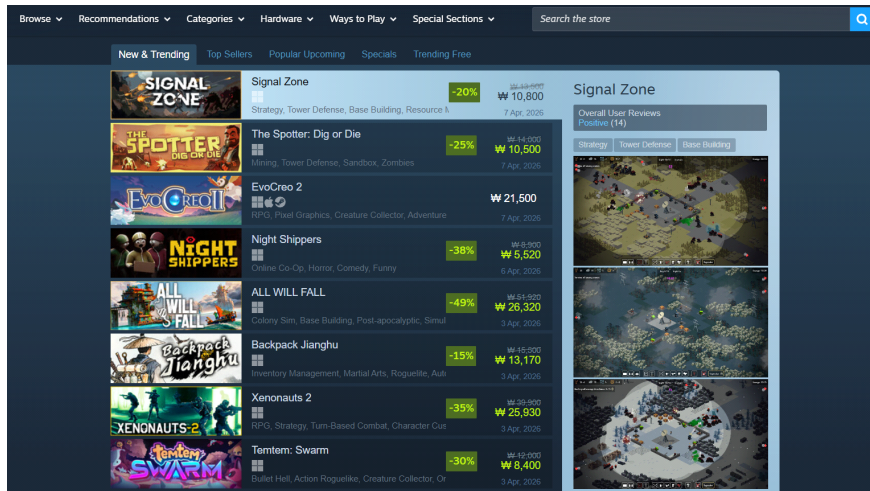
1) 저작물의 ‘공중전달(communication to the public)’을 저작권자의 배타적 권리가 미치는 제한행위로 규정

2) ‘복제(copying)’를 모든 유형의 저작물에 대한 제한행위로 규정

3) Chris Cooke, “PRS sues gaming platform giant Steam for copyright infringement”, Complete Music Update, 2026.03.09., <https://completemusicupdate.com/prs-sues-gaming-platform-giant-steam-for-copyright-infringement/>

게임을 구매하고 개발자들이 경쟁하는 게임 마켓플레이스의 표준<sup>4)</sup>이라고 평가하였다. 이용자들은 스팀에서 게임을 구매·다운로드하는 것 이외에도, 모바일 기기로 게임을 스트리밍하거나 실시간 게임플레이를 방송할 수 있다.

[그림] 스팀 메인페이지



출처: 스팀 공식 홈페이지(<https://store.steampowered.com/>)

이처럼 스팀이 게임 산업 내 핵심적인 플랫폼으로 자리매김하고 있음에도 불구하고, PRS는 밸브가 2003년 스팀 서비스를 출시한 이후 약 23년간 단 한 차례도 CDPA 체계에 부합하는 라이선스를 취득하지 않았다고 설명하였다.

• 여타 게임 유통 플랫폼의 PRS 라이선스 체결 현황

PRS는 과거 소니 컴퓨터 엔터테인먼트 유럽(Sony Computer Entertainment Europe)과 유럽 플레이스테이션 스토어(PlayStation Store) 게임 음악 공연권에 대한 라이선스 계약을 체결한 바 있다. 당시 소니는 PRS측에 2007년 2월부터 2013년 12월까지 유럽 플레이스테이션 스토어에서 다운로드된 게임 목록을 제공하였으며, PRS는 해당 목록을 바탕으로 게임 내 사용된 음악의 작곡가·퍼블리셔에게 저작권료를 지급하였다. 또한, 이후에도 PRS는 유사한 방식으로 2020~2021년 간 영국, 아일랜드, 프랑스, 독일 등 유럽 23개국에서 다운로드된 게임을 검토하고, 음악의 사용 실적을 기준으로 권리자에게 저작권료를 지급하였다.

마이크로소프트(Microsoft) 역시 자사의 엑스박스(Xbox) 플랫폼에 대해 PRS 체계에 포함되는 라이선스인 GEOL(General Entertainment Online Licence)을 취득한 상태이다. GEOL은 온라인 시청각 콘텐츠 제공 사업자를 대상으로 발급되는 라이선스로, 음악 사용 시간, 이용자 수, 음악 활용 수준 등을 고려하여 개별적으로 라이선스 비용을 산정한다.

4) Chris Cooke, "PRS sues gaming platform giant Steam for copyright infringement", Complete Music Update, 2026.03.09., <https://completemusicupdate.com/prs-sues-gaming-platform-giant-steam-for-copyright-infringement/>



[표] 소니와 마이크로소프트의 PRS 라이선스 비교

구분	소니	마이크로소프트
라이선스 유형	플레이스테이션 스토어 게임 음악 공연권 라이선스	GEOL (General Entertainment Online Licence)
라이선스 체결 플랫폼	유럽 플레이스테이션 스토어 (Europe Playstation Store)	엑스박스(Xbox)
대상 권리	게임이 온라인으로 제공될 때 발생하는 공연권	공연권 및 복제권 공동 라이선스
적용 범위	플레이스테이션 스토어를 통해 다운로드되는 게임 내 음악	온라인 시청각 콘텐츠 제공 서비스 전반

출처: PRS 소니 라이선스 관련 공식 문서, MCPS(PRS) 온라인 라이선스 정책 자료를 바탕으로 재구성

• 음악 권리단체-게임·디지털 플랫폼 간 분쟁 선례

이번 PRS의 벨브 소송 사건 이전에도, 타국에서 음악 권리단체가 게임 및 디지털 플랫폼을 상대로 저작권 라이선스를 요구하는 사례가 확인된 바 있다. 대표적으로, 미국의 음악출판사협회(National Music Publishers' Association, 이하 NMPA)는 지난 2021년 6월 게임 플랫폼 로블록스(Roblox)의 이용자들이 저작권이 있는 음악을 무단으로 업로드·공유하고 있으며, 로블록스가 이를 통해 수익을 얻고 있다는 이유로 소송을 제기하고 2억 달러(원화 약 3,060억 원)<sup>5)</sup> 규모의 손해배상을 청구하였다. 해당 사건은 같은 해 9월 양측 간 합의로 마무리되었으며, 주요 합의 내용에는 옵트인(opt-in)\* 라이선스 체계를 마련하는 방안이 포함되었다.

\* 옵트인(opt-in): 권리가자 스스로 참여 의사를 밝혀 라이선스 체계에 들어가는 방식

아울러, 같은 시기 NMPA는 게임 라이브스트리밍 플랫폼 트위치(Twitch)와도 관련 협약을 체결하여 음악 권리가자 자신의 음악을 플랫폼에 라이선스할 수 있는 경로를 마련하였다. 이러한 일련의 사례는 다양한 음악 권리단체들이 디지털 플랫폼 내 라이선스 체계를 확립하기 위해 노력하고 있다는 점을 시사하며, 이번 PRS의 벨브 소송 사건 역시 이러한 패러다임의 연장선에 있는 것으로 분석된다.

III. 결론 : 게임 유통 플랫폼 운영사들의 음악 저작권 관리 과제

• 게임 유통 플랫폼 사업자의 저작권 책임 범위 확대 가능성

통상 게임 내 음악 관련 저작권 문제는 게임 개발사 또는 퍼블리셔가 싱크 라이선스를 취득하는 것으로 해결된다는 인식이 일반적이었으나, 이번 PRS의 벨브 소송 사례는 게임 유통 플랫폼 운영사 측에도 별도의 라이선스 의무가 부과될 수 있다는 점을 시사하였다. 이는 디지털 콘텐츠 유통 구조가 고도화됨에 따라, 저작권의 각 권리 유형별 책임 부담 주체 역시 세분화되고 있음을 의미한다.

5) 1달러=1,530.50원(2026.04.01, KEB 하나은행 매매기준율 적용, 이하 동일)

게임 유통 플랫폼 운영사에 음악 저작권 라이선스 의무가 부과될 경우, 게임 유통 시장의 수익 구조에도 변화가 발생할 수 있을 것으로 보인다. 현재 스팀에는 약 12만 개 이상의 게임이 등록되어 있으며, 이 가운데 상당수가 음악을 포함하고 있다. 플랫폼 운영사가 개별 게임 내 삽입된 음악의 저작권 현황을 파악하고 이에 대한 라이선스 비용을 부담해야 한다면, 이는 유통 수수료 인상이나 게임 판매 가격 조정으로 이어질 수 있다.

밸브 측은 아직까지 PRS 소송에 대한 공식적인 입장을 발표하지 않고 있으나, 이번 PRS의 밸브 소송 사례는 글로벌 게임 유통 플랫폼들이 저작권 책임에 대한 인식을 재고하는 계기가 될 것으로 전망된다. 아울러, 이번 소송의 결과는 향후 게임 유통 플랫폼뿐 아니라 음악이 포함된 디지털 콘텐츠를 유통하는 다양한 플랫폼 사업자들이 참고할 수 있는 선례를 남길 것으로 전망된다.

### 참고문헌

- PRS for Music, “Video gaming platform Steam faces legal proceedings”, PRS for Music, 2026.03.09., <https://www.prsformusic.com/press/2026/video-gaming-platform-steam-faces-legal-proceedings>
- Beth Simpson, “PRS sues Steam for using music without a licence – for 23 years”, musicradar, 2026.03.11., <https://www.musicradar.com/music-industry/prs-sues-steam-for-using-music-without-a-licence-for-23-years>
- Vikki Blake, “Valve sued by The Performing Right Society for allegedly using its members' musical works "without permission"”, Games Industry.biz, 2026.03.09., <https://www.gamesindustry.biz/valve-sued-by-the-performing-right-society-for-using-its-members-musical-works-without-permission>
- Chris Cooke, “PRS sues gaming platform giant Steam for copyright infringement”, Complete Music Update, 2026.03.09., <https://completemusicupdate.com/prs-sues-gaming-platform-giant-steam-for-copyright-infringement/>
- MCPS, “Online Licensing”, MCPS, 2026.04.08. 접속 기준, <https://www.prsformusic.com/-/media/files/prs-for-music/royalties/mcps-updates/5-online-licensing-policyupdated062024>
- IMRO, “Sony PlayStation Game Music Royalties”, IMRO, 2022.08.17., <https://imro.ie/featured-article/sony-playstation-game-music-royalties-2/>
- Eric T. Straka, “Music Publishers Reach Settlements with Gaming Platforms Regarding the Use of Copyrighted Music”, Winston & Strawn, 2021.11.09., <https://www.winston.com/en/blogs-and-podcasts/the-playbook/music-publishers-reach-settlements-with-gaming-platforms-regarding-the-use-of-copyrighted-music>
- Legislation.gov.uk, “Copyright, Designs and Patents Act 1988”, Legislation.gov.uk, 2026.04.21. 접속 기준, <https://www.legislation.gov.uk/ukpga/1988/48>

### 기술용어

순번	용어	설명
1	싱크 라이선스 (sync license)	저작권이 있는 음악을 영상물이나 게임 등 시청각물과 동기화하여 삽입·이용할 수 있도록 허락하는 권리처리 방식
2	옵트인 (opt-in)	권리자가 스스로 참여 의사를 밝혀 라이선싱 체계에 들어가는 방식



# 마이크로소프트, 음성·이미지 생성에 특화된 자체 AI 모델 3종 공개

## 뉴스 브리프

2026년 4월 마이크로소프트는 음성 인식, 음성 합성, 이미지 생성에 특화된 자체 AI 모델인 MAI 3종을 공개했다. 이는 오픈AI와의 협력을 유지하면서도 자체 AI 모델 역량을 병행 강화하려는 움직임으로, AI 모델 시장에서 협력과 자립을 동시에 추구하는 전략적 선택이 확산되고 있음을 보여준다. 본 보고서는 마이크로소프트가 공개한 MAI 모델 3종의 기술적 특징과 활용 가능성을 분석하고, 자체 AI 모델 확대가 오픈AI와의 협력 구도에 미치는 영향을 살펴본다.

## 뉴스 플러스

### I. 서론: 마이크로소프트의 자체 AI 모델 공개

#### • 마이크로소프트의 오픈AI 협력 유지 속 자체 AI 모델인 MAI 3종 공개

2026년 4월 2일, 마이크로소프트는 음성 인식, 음성 합성, 이미지 생성에 특화된 자체 AI 모델 3종의 공개 베타 버전을 발표했다. 이번에 공개된 모델은 MAI-Transcribe-1, MAI-Voice-1, MAI-Image-2로, 모두 마이크로소프트 파운드리(Microsoft Foundry)를 통해 제공된다. 마이크로소프트는 2019년부터 오픈AI(OpenAI)와 협력해왔으며, 2025년 10월 체결한 새 계약에서도 오픈AI를 자사의 핵심 프론티어 모델 파트너로 유지한다고 밝힌 바 있다. 또한 2026년 2월 양사 공동성명에서도 연구, 엔지니어링, 제품 개발 전반에서 협력이 계속되고 있음을 재확인했다. 이런 점에서 이번 발표는 기존 협력 관계의 종료가 아닌, 오픈AI와의 협력을 유지하는 동시에 자체 AI 모델 개발도 병행·확대하려는 움직임으로 해석된다.

미국 IT 전문매체 더 레지스터(The Register)는 이번 발표를 두고, 마이크로소프트가 오픈AI의 투자자가 아닌 직접 경쟁자에 가까워 보이게 하는 움직임이라고 평가했다.<sup>1)</sup> 실제로 오픈AI 역시 음성 인식, 음성 생성, 텍스트-이미지 생성 모델을 제공하고 있어, 양사는 동일한 시장에서 경쟁 가능한 제품군을 점차

1) Thomas Claburn, "Microsoft shivs OpenAI with three new AI models for speech and images", The Register, 2026.04.02., [https://www.theregister.com/2026/04/02/microsoft\\_models\\_homegrown\\_ai\\_models/](https://www.theregister.com/2026/04/02/microsoft_models_homegrown_ai_models/)

갖춰가고 있다. 다만 이번 발표는 어디까지나 베타 버전 모델이라는 점에서, 완전한 상용 안정화보다는 기술 공개와 생태계 확장의 성격이 더 크다고 볼 수 있다.

또한 이번 발표는 단순한 연구 성과 공개에 그치지 않는다는 점에서도 의미가 있다. 마이크로소프트는 이번 모델들이 이미 코파일럿(Copilot), Bing(Bing), 파워포인트(PowerPoint), 애저 스피치(Azure Speech) 등 자사 제품군에 적용되고 있다고 설명했다. 이는 마이크로소프트가 자체 AI 모델을 단순히 보유하는 수준을 넘어, 자사 서비스 전반에 실제로 적용하고 있음을 보여준다. 다만 현재 이들 모델은 공개 미리 보기 단계로 제공되고 있어, 일부 기능과 운영 환경에는 제약이 존재한다는 점도 함께 고려할 필요가 있다.

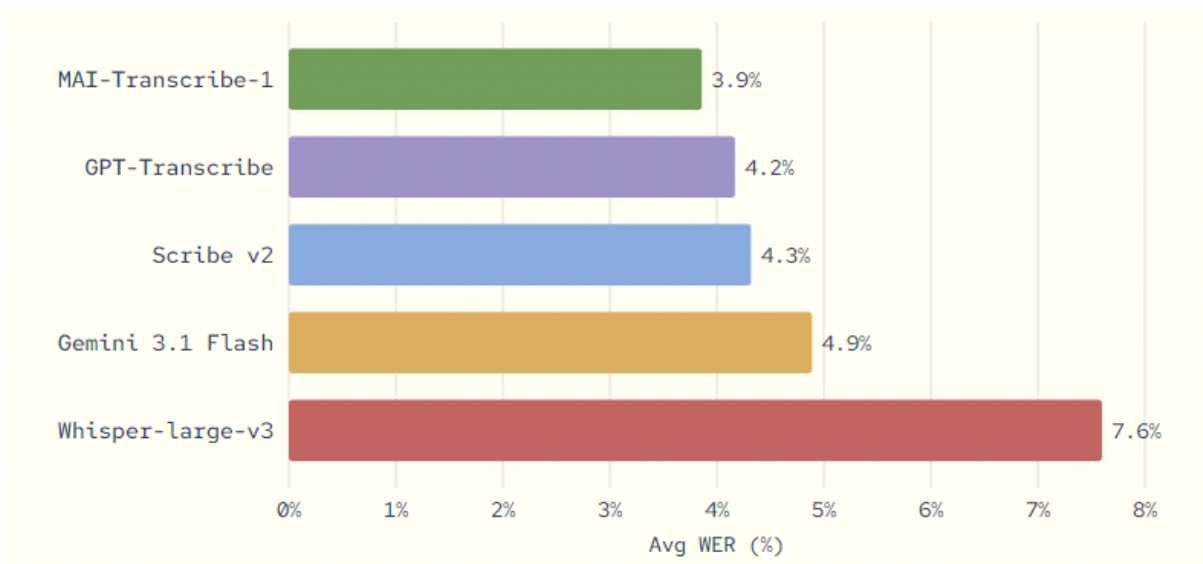
\* 마이크로소프트 파운드리(Microsoft Foundry): AI 모델, 개발 도구, 인프라, 배포-운영 기능을 통합 제공하는 마이크로소프트의 기업용 AI 플랫폼

## II. 본론: MAI 모델 3종의 기술적 특징과 활용 가능성

### • MAI-Transcribe-1의 다국어 음성 인식 성능과 비용 효율성

MAI-Transcribe-1은 높은 정확도와 효율성을 동시에 지향하는 음성 인식 모델이다. 마이크로소프트에 따르면, 이 모델은 FLEURS 벤치마크\*에서 25개 언어를 대상으로 평균 단어 오류율(Word Error Rate, WER)\*\* 3.9%를 기록했다. 이는 비교 대상으로 제시된 GPT-Transcribe 4.2%, Scribe v2 4.3%, Gemini 3.1 Flash 4.9%, Whisper-large-v3 7.6%보다 낮은 수치로, 높은 정확도임을 의미한다. 마이크로소프트는 MAI-Transcribe-1이 25개 언어 중 11개 핵심 언어에서 1위를 기록했으며, 나머지 언어들 가운데서도 다수 영역에서 경쟁 모델 대비 우수한 성능을 보였다고 설명했다.

[그림 1] MAI-Transcribe-1과 주요 음성 인식 모델의 평균 단어 오류율 비교



출처: Mustafa Suleyman, "Today we're announcing 3 new world class MAI models, available in Foundry", Microsoft AI, 2026.04.02., <https://microsoft.ai/news/today-were-announcing-3-new-world-class-mai-models-available-in-foundry/>

MAI-Transcribe-1은 효율성 측면에서도 강점을 보인다. 마이크로소프트는 MAI-Transcribe-1이 다른 음성 전사(transcription)<sup>\*\*\*</sup> 모델과 비교해 약 50% 낮은 GPU 비용 수준에서 경쟁력 있는 정확도를 제공하며, 저장된 오디오 파일을 일괄 처리하는 배치 전사(batch transcription)<sup>\*\*\*\*</sup> 속도도 기존 Azure Fast 방식보다 2.5배 빠르다고 밝혔다. 가격은 시간당 0.36달러 (원화 약 550원)<sup>2)</sup>부터 시작하며, WAV, MP3, FLAC 형식의 300MB 미만 오디오 파일을 지원한다.

적용 분야로는 대규모 행사 및 회의의 실시간 자막 제공, 미디어 자막 제작과 아카이빙, 교육 및 훈련 콘텐츠를 텍스트로 기록·정리하는 작업, 고객 인터뷰 및 포커스 그룹 분석 등이 제시된다. 또한 가상 비서와 콜센터 워크플로에 실시간 전사 기능을 적용해 음성 기반 인터페이스와 후속 요약 기능을 지원할 수 있다. 마이크로소프트는 코파일럿의 음성 모드에서 발화를 텍스트로 변환하는 기능과 새로운 받아쓰기 기능에 MAI-Transcribe-1이 활용되고 있다고 밝혔다.

\* FLEURS 벤치마크: 여러 언어에서 음성 인식 성능을 비교하기 위해 쓰이는 대표 평가 기준

\*\* 단어 오류율(Word Error Rate, WER): 음성을 글로 바꿨을 때 틀린 단어가 얼마나 되는지를 보여주는 지표로, 낮을수록 정확도가 높음

\*\*\* 전사(transcription): 음성이나 발화를 문자 텍스트로 옮기는 작업 또는 기능.

\*\*\*\* 배치 전사(batch transcription): 실시간 대화가 아니라, 이미 저장된 오디오 파일을 한꺼번에 불러와 텍스트로 변환하는 방식

## • MAI-Voice-1의 자연스러운 음성 생성과 감정 표현 제어 기능

MAI-Voice-1은 마이크로소프트가 자체 음성 기반 기술을 토대로 개발한 음성 합성 모델이다. 이 모델은 입력된 텍스트의 전체 맥락을 반영해 톤과 감정, 말하기 방식을 조정함으로써 보다 자연스럽고 대화형에 가까운 음성을 생성하는 데 초점을 맞추고 있다. 단순히 문장을 기계적으로 읽는 수준을 넘어, 감정 표현과 말투의 변화를 반영하면서도 긴 형식의 콘텐츠에서 일관된 음성 특성을 유지할 수 있도록 설계된 것이 특징이다.

MAI-Voice-1은 속도와 효율성이 주요 강점으로 제시된다. 이 모델은 이론적으로 단일 GPU에서 1초 내에 60초 분량의 오디오를 생성할 수 있다. 가격은 100만 문자당 22달러(원화 약 33,671원)부터 제공한다. 개발자는 SSML(Speech Synthesis Markup Language)<sup>\*</sup>을 활용하여 말하기 스타일에 영향을 줄 수 있으며, 기쁨, 공감, 흥분 등 감정 표현을 보다 세밀하게 조정할 수 있다.

MAI-Voice-1은 개인 음성 복제 기능도 지원한다. 이는 단 10초 분량의 오디오 샘플만으로 특정 화자의 음성을 복제할 수 있는 기능이다. 다만 이 기능은 마이크로소프트의 책임 있는 AI 정책에 따라 제한된 접근 심사를 통해서만 사용할 수 있다. 현재 공개 미리 보기 기준으로는 영어(미국) 음성 6종이 제공되며, 대화형 AI, 고객 서비스, 창의적 콘텐츠 제작, 장문 내레이션 등 다양한 활용이 가능하다.

\* SSML(Speech Synthesis Markup Language): 음성 합성 시 속도, 높낮이, 감정 표현, 말투 등을 조정할 수 있도록 만든 마크업 언어

### • MAI-Image-2의 고품질 이미지 생성 성능과 활용 범위

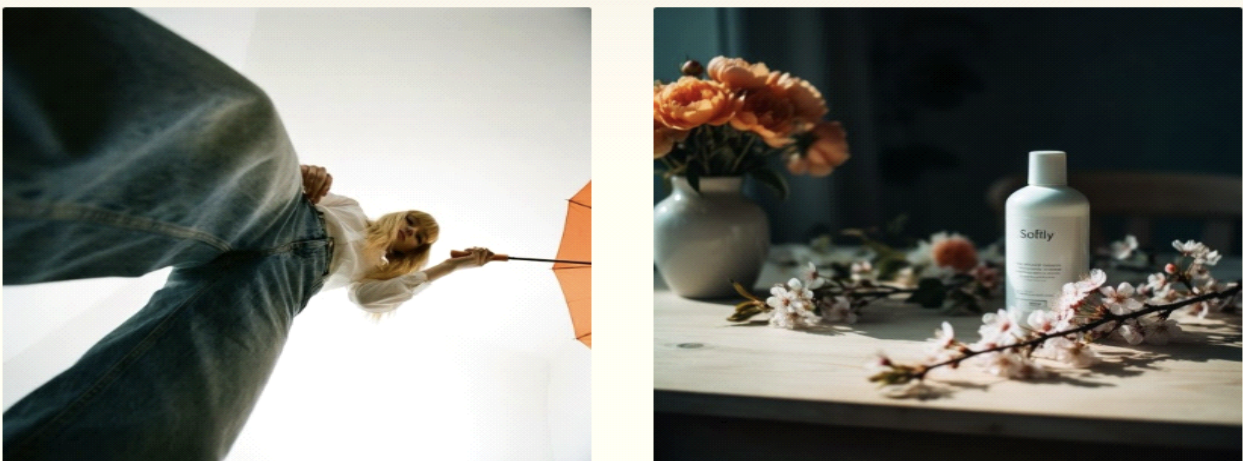
MAI-Image-2는 자연어 프롬프트\*를 바탕으로 고품질 이미지를 생성하도록 설계된 텍스트-이미지 생성 모델이다. 이 모델은 확산 기반 생성 방식(Diffusion-based generation)\*\*을 적용해 입력된 설명에 맞는 이미지를 점진적으로 정교하게 만들어 간다. 이를 통해 개념 시각화, 크리에이티브 콘텐츠 제작, 이미지 디자인 업무 등 다양한 창의적 작업에 활용될 수 있다.

마이크로소프트는 MAI-Image-2가 사진작가, 디자이너, 시각 스토리텔러와의 협업을 바탕으로 개발됐으며, 자연스러운 조명과 피부 톤, 질감 표현, 이미지 내 텍스트 렌더링 등 실무에서 중요한 요소를 중점적으로 개선했다고 설명했다. 또한 Arena.ai 리더보드\*\*\*에서 텍스트-이미지 모델 패밀리 3위에 올랐고, 실제 운영 데이터 기준으로 마이크로소프트 파운드리와 코파일럿에서 최소 2배 빠른 생성 속도를 기록했다고 밝혔다.

입력과 출력 조건도 구체적으로 제시돼 있다. MAI-Image-2는 최대 3만 2,000개 토큰의 텍스트 입력을 받아 하나의 PNG 이미지를 생성한다. 이미지 크기는 너비와 높이가 각각 최소 768픽셀 이상이어야 하며, 전체 픽셀 수는 1,048,576을 초과할 수 없다. 가격은 텍스트 입력 100만 토큰당 5달러(원화 약 7,652원), 이미지 출력 100만 토큰당 33달러(원화 약 50,506원)부터 시작한다.

활용 분야로는 미디어와 크리에이티브 아이디어 도출, 기업 커뮤니케이션과 내부 브랜딩, UX 및 제품 콘셉트 시각화 등이 제시된다. 마이크로소프트는 세계적 마케팅 기업 WPP를 초기 활용 사례로 소개했으며, 코파일럿, Bing 이미지 크리에이터, 파워포인트에도 MAI-Image-2를 단계적으로 적용하고 있다고 밝혔다.

[그림 2] MAI-Image-2 기반 이미지 생성 예시



출처: Mustafa Suleyman, "Today we're announcing 3 new world class MAI models, available in Foundry", Microsoft AI, 2026.04.02., <https://microsoft.ai/news/today-were-announcing-3-new-world-class-mai-models-available-in-foundry/>

\* 자연어 프롬프트: 사용자가 일상적인 문장으로 입력하는 지시문이나 설명

\*\* 확산 기반 생성 방식(Diffusion-based generation): 무작위에 가까운 상태에서 시작해 이미지를 단계적으로 정제하면서 결과물을 만들어 가는 생성 방식

\*\*\* Arena.ai 리더보드: 여러 이미지 생성 모델의 성능을 비교·평가해 순위를 제시하는 지표

### III. 결론 및 전망: AI 모델 시장의 변화와 향후 과제

#### • 마이크로소프트의 협력 유지 속 자체 AI 모델 확대와 시장 경쟁

마이크로소프트의 자체 AI 모델 3종 공개는 오픈AI와의 협력을 유지하면서도 자체 AI 모델 역량을 병행 강화하려는 움직임으로 볼 수 있다. 이번 MAI 모델 발표는 마이크로소프트가 음성 인식, 음성 합성, 이미지 생성 영역에서 오픈AI 모델과 함께 활용할 수 있는 자체 AI 모델 역량을 확대하고 있음을 보여주는 사례로 해석될 수 있다. 이는 단일 공급자 의존 리스크를 분산하고, 모델 성능과 비용을 보다 유연하게 관리할 수 있는 기반을 넓혀가는 움직임으로 볼 수 있다.

향후 AI 모델 시장은 기술 자립과 생태계 주도권을 둘러싼 경쟁이 점차 심화될 가능성이 있다. 이러한 흐름 속에서 마이크로소프트도 2026년 3월 조직 개편을 통해 코파일럿 조직을 재정비하고 자체 AI 모델 확대에 맞는 체계를 갖춰가고 있다. 다만 이러한 변화가 곧바로 오픈AI와의 결별을 의미하는 것은 아니다. 마이크로소프트는 2032년까지 파트너십을 유지하기로 했으며, 오픈AI 모델 역시 마이크로소프트 파운드리를 통해 계속 제공될 것으로 보인다. 결국 마이크로소프트의 이번 행보는 특정 파트너에 대한 의존을 줄이면서도 협력 관계를 유지하는, 이른바 '협력 속 자립' 전략이 AI 모델 시장에서 하나의 경쟁 방식으로 자리 잡아 가고 있음을 보여준다.

#### 참고문헌

- Thomas Claburn, "Microsoft shivs OpenAI with three new AI models for speech and images", The Register, 2026.04.02., [https://www.theregister.com/2026/04/02/microsoft\\_models\\_homegrown\\_ai\\_models/](https://www.theregister.com/2026/04/02/microsoft_models_homegrown_ai_models/)
- Mustafa Suleyman, "Today we're announcing 3 new world class MAI models, available in Foundry", Microsoft AI, 2026.04.02., <https://microsoft.ai/news/today-were-announcing-3-new-world-class-mai-models-available-in-foundry/>
- Naomi Money Penny, "Introducing MAI-Transcribe-1, MAI-Voice-1, and MAI-Image-2 in Microsoft Foundry", Microsoft Foundry Blog, 2026.04.02., <https://techcommunity.microsoft.com/blog/azure-ai-foundry-blog/introducing-mai-transcribe-1-mai-voice-1-and-mai-image-2-in-microsoft-foundry/4507787>
- Microsoft Learn, "Azure Speech의 MAI-Transcribe-1(미리 보기)", 2026.04.07. 접속 기준, <https://learn.microsoft.com/ko-kr/azure/ai-services/speech-service/mai-transcribe>
- Microsoft Learn, "Azure Speech의 MAI-Voice-1(미리 보기)", 2026.04.07. 접속 기준, <https://learn.microsoft.com/ko-kr/azure/ai-services/speech-service/mai-voices>
- Microsoft Learn, "Microsoft Foundry에서 MAI-Image-2 배포 및 사용(미리 보기)", 2026.04.07. 접속 기준, <https://learn.microsoft.com/ko-kr/azure/foundry/foundry-models/how-to/use-foundry-models-mai>



## 기술용어

순번	용어	설명
1	마이크로소프트 파운드리 (Microsoft Foundry)	AI 모델, 개발 도구, 인프라, 배포·운영 기능을 통합 제공하는 마이크로소프트의 기업용 AI 플랫폼
2	FLEURS 벤치마크	여러 언어에서 음성 인식 성능을 비교하기 위해 쓰이는 대표 평가 기준
3	단어 오류율 (Word Error Rate, WER)	음성을 글로 바꿨을 때 틀린 단어가 얼마나 되는지를 보여주는 지표로, 낮을수록 정확도가 높음
4	전사 (transcription)	음성이나 발화를 문자 텍스트로 옮기는 작업 또는 기능
5	배치 전사 (batch transcription)	실시간 대화가 아니라, 이미 저장된 오디오 파일을 한꺼번에 불러와 텍스트로 변환하는 방식
6	SSML (Speech Synthesis Markup Language)	음성 합성 시 속도, 높낮이, 감정 표현, 말투 등을 조정할 수 있도록 만든 마크업 언어
7	자연어 프롬프트	사용자가 일상적인 문장으로 입력하는 지시문이나 설명
8	확산 기반 생성 방식 (Diffusion-based generation)	무작위에 가까운 상태에서 시작해 이미지를 단계적으로 정제하면서 결과물을 만들어 가는 생성 방식
9	Arena.ai 리더보드	여러 이미지 생성 모델의 성능을 비교·평가해 순위를 제시하는 지표