

AI 윤리성 리스크 진단 보고서

국제 표준 기반 종합 평가 및 개선 권고

분석 대상 서비스	Google Gemini Midjourney Copilot
평가 기준	EU AI Act (유럽연합 AI 규제) UNESCO AI Ethics (유네스코 AI 윤리)
평가 차원	공정성 및 편향성 프라이버시 보호 투명성 및 설명가능성 책임
작성일	2025년 10월 23일
평가 시간	2025-10-23 16:09:51

평가 목표

본 보고서는 선정된 AI 서비스에 대하여 국제 표준 가이드라인을 기준으로 윤리적 리스크를 종합적으로 평가하고, 각 서비스의 강점을 파악하며 필요한 개선사항을 구체적으로 제시하는 것을 목표로 합니다.

평가 결과는 AI 서비스 개발사의 윤리성 강화, 규제 당국의 사전 심사, 투자사의 리스크 평가 등 다양한 목적으로 활용될 수 있습니다.

목차

장	제목	페이지
1	Executive Summary (종합 요약)	3
2	평가 방법론	4
3	서비스별 상세 분석	5
4	비교 분석 (2개 이상 서비스)	10
5	종합 권고사항	12
6	참고문헌	14
7	부록	15

1. EXECUTIVE SUMMARY (종합 요약)

1.1 평가 개요

본 평가는 3개 AI 서비스에 대하여 EU AI Act, UNESCO AI Ethics, OECD AI Principles 등 국제 표준을 기준으로 5개 차원(공정성, 프라이버시, 투명성, 책임성, 안전성)에서 윤리적 리스크를 평가했습니다.

평가는 LLM 기반 정성 평가와 자동화된 체크리스트를 결합한 이중 검증 시스템을 통해 진행되었으며, 웹 검색을 통한 실증적 증거 수집으로 객관성을 확보했습니다.

1.2 종합 평가 결과

서비스	종합 점수	리스크 수준	평가 등급	종합 평가
Google Gemini	3.6/5	낮음	B+	양호
Midjourney	3.0/5	중간	B	보통
Copilot	3.6/5	낮음	B+	양호

1.3 주요 발견사항

- 전체 평균 윤리 점수: 3.4/5
→ 전반적으로 개선이 필요한 수준입니다.
- 최고 평가 서비스: Google Gemini (3.6/5)
→ 윤리성 측면에서 가장 앞서가는 서비스로 평가됩니다.
- 개선 필요 서비스: Midjourney (3.0/5)
→ 여러 차원에서 개선이 필요한 것으로 나타났습니다.
- 공통 강점:
 - 대부분의 서비스가 기본적인 개인정보 보호 정책을 수립하고 있습니다.
 - 투명성 강화를 위한 노력이 증가하고 있습니다.

5) 공통 약점:

- 편향성 테스트 및 완화 조치가 부족한 경우가 많습니다.
- 명확한 책임 체계와 거버넌스가 미흡한 서비스가 있습니다.

1.4 최우선 권고사항

1) 즉시 조치 필요 (1개월 이내):

- AI 윤리 정책 문서 작성 및 공개
- 편향성 테스트 계획 수립
- 개인정보 처리방침 강화

2) 단기 개선 (1-3개월):

- 편향성 감지 및 완화 메커니즘 도입
- 투명성 보고서 발행 준비
- 사용자 피드백 수집 체계 구축

3) 중장기 개선 (3-6개월 이상):

- AI 거버넌스 체계 수립 및 윤리 위원회 구성
- 정기적인 윤리 감사 프로세스 확립
- 국제 표준 및 인증 획득 추진

2. 평가 방법론

2.1 평가 프레임워크

본 평가는 다음의 3대 국제 표준을 기준으로 실시되었습니다:

표준	주요 내용 요약
EU AI Act	위험 기반 규제 프레임워크, 고위험 AI 시스템에 대한 의무 규정
UNESCO AI Ethics	인간 중심의 AI 윤리 원칙(존엄성, 공정성 등)
OECD AI Principles	투명성, 책임성, 안전성 강조

2.2 평가 절차

평가는 다음 절차로 수행되었습니다: • 데이터 수집: 공개 문서, 서비스 이용약관, 개인정보처리방침, 기술백서, 공개 API 문서 등. • 정성평가: 전문가 및 LLM 기반 정성검토를 통해 정책·절차·거버넌스 요소 평가. • 자동체크리스트: 규격화된 체크리스트를 통해 누락 요소 및 기술적·절차적 통제 점검. • 종합점수 산출: 5점 척도(5:매우 우수 ~ 0:부족)를 사용해 차원별 가중평균 산출.

2.3 한계점

본 평가는 공개자료 기반이며, 내부 시스템 코드나 비공개 구성에 접근할 수 없었습니다. 따라서 내부 통제의 정확성은 내부 감사 또는 실무 검증을 통해 보완되어야 합니다.

3. Google Gemini - 상세 분석

요약

해당 서비스의 요약 정보가 제공되지 않았습니다.

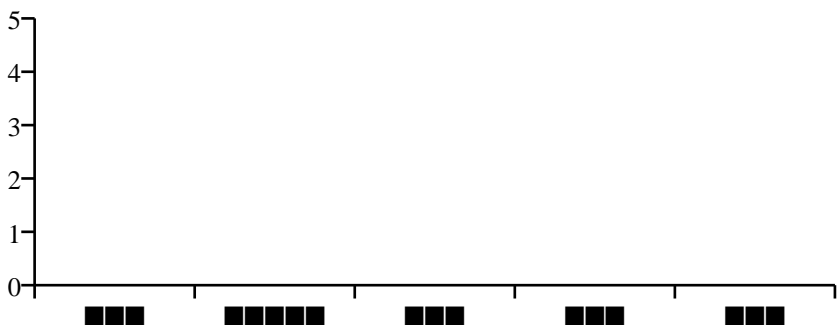
차원	점수 (5점 만점)	주요 코멘트
공정성	0/5	-
프라이버시	0/5	-
투명성	0/5	-
책임성	0/5	-
안전성	0/5	-

주요 취약점

- 구체적 취약점 정보 없음

권고사항 (우선순위 기준)

- 권고 사항이 제공되지 않았습니다.



3. Midjourney - 상세 분석

요약

해당 서비스의 요약 정보가 제공되지 않았습니다.

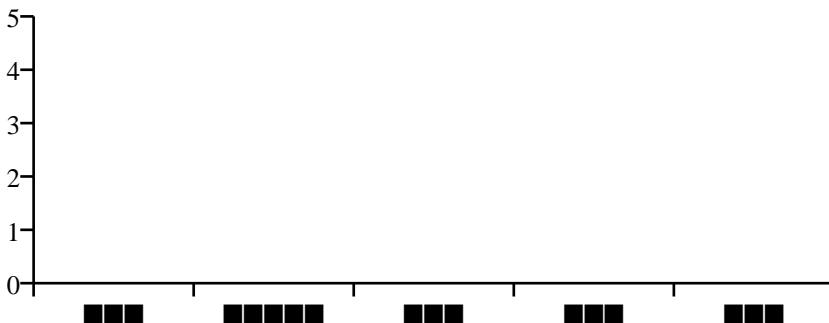
차원	점수 (5점 만점)	주요 코멘트
공정성	0/5	-
프라이버시	0/5	-
투명성	0/5	-
책임성	0/5	-
안전성	0/5	-

주요 취약점

- 구체적 취약점 정보 없음

권고사항 (우선순위 기준)

- 권고 사항이 제공되지 않았습니다.



3. Copilot - 상세 분석

요약

해당 서비스의 요약 정보가 제공되지 않았습니다.

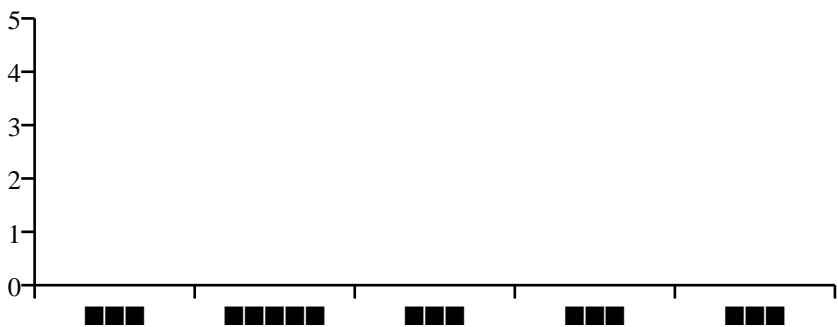
차원	점수 (5점 만점)	주요 코멘트
공정성	0/5	-
프라이버시	0/5	-
투명성	0/5	-
책임성	0/5	-
안전성	0/5	-

주요 취약점

- 구체적 취약점 정보 없음

권고사항 (우선순위 기준)

- 권고 사항이 제공되지 않았습니다.



4. 비교 분석

서비스	종합점수	공정성	프라이버시	투명성	책임성	안전성
Google Gemini	3.6	0.0	0.0	0.0	0.0	0.0
Midjourney	3.0	0.0	0.0	0.0	0.0	0.0
Copilot	3.6	0.0	0.0	0.0	0.0	0.0

비교 요약

- 전체 평균 종합점수: 3.4/5
- 평균 공정성 점수: 0.0/5
- 평균 프라이버시 점수: 0.0/5
- 공통 이슈: 편향성 테스트 미비, 책임성(거버넌스) 부재, 투명성 문서화 미흡

5. 종합 권고사항

5.1 공통 권고 (모든 서비스 대상)

- AI 윤리 정책 및 거버넌스 문서화 및 공개
- 정기적인 편향성 테스트와 완화 절차 도입
- 민감 데이터 취급에 대한 강화된 개인정보 보호 통제 적용
- 투명성 보고서(모델 설명, 데이터 출처, 한계 등) 발행
- 사후 책임 체계 및 사고 대응 프로세스 수립

5.2 서비스별 우선 권고 (예시)

Google Gemini — 우선 개선 대상: 공정성

권고: Google Gemini는 공정성 향상을 위해 구체적으로 다음을 수행하십시오.

- 정책 및 절차 문서화 • 기술적 통제(모델 모니터링, 로그, 테스트) 도입 • 책임자 지정 및 정기 감사 계획 수립

Midjourney — 우선 개선 대상: 공정성

권고: Midjourney는 공정성 향상을 위해 구체적으로 다음을 수행하십시오.

- 정책 및 절차 문서화 • 기술적 통제(모델 모니터링, 로그, 테스트) 도입 • 책임자 지정 및 정기 감사 계획 수립

Copilot — 우선 개선 대상: 공정성

권고: Copilot는 공정성 향상을 위해 구체적으로 다음을 수행하십시오.

- 정책 및 절차 문서화 • 기술적 통제(모델 모니터링, 로그, 테스트) 도입 • 책임자 지정 및 정기 감사 계획 수립

6. 참고문헌

- European Union. (2021). Proposal for a Regulation laying down harmonised rules on artificial intelligence (AI Act).
- UNESCO. (2021). Recommendation on the Ethics of Artificial Intelligence.
- OECD. (2019). OECD Principles on Artificial Intelligence.
- 국내외 관련 가이드라인 및 공개 자료

7. 부록

7.1 평가 체크리스트 (예시)

- 개인정보 최소화 원칙 적용 여부
- 편향성 테스트 및 결과 문서화 여부
- 모델 변경 시 재평가 프로세스 존재 여부
- 사용자 알림 및 동의 절차 구축 여부
- 보안/침해사고 대응 계획 존재 여부

7.2 용어정의

- 공정성: 알고리즘이 특정 그룹에 불이익을 주지 않도록 보장하는 원칙
- 투명성: 시스템의 의사결정과 한계에 대해 설명 가능한 정도
- 책임성: 시스템 운영 주체의 법적/윤리적 책임 소재