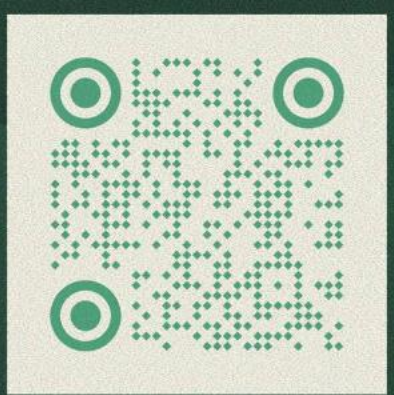
AI시민사회포럼

폭주하는 AI를 멈춰라

AI의 공공성과 책임성을 묻다



자료집 다운로드

후원 : 아름다운재단

주최 : 인공지능 책임성과 공공성 강화를 위한 시민사회 공동행동

AI시민사회포럼 프로그램 안내

13:00~13:10 개회식

13:10~14:10 강좌

AI 기술 동향 (조경숙, 도토리랩스 대표)

AI 정책 동향 (최호웅, 민변 디지털정보위원회 위원장 / 변호사)

14:10~14:30 휴식

14:30~16:00 세션

* 트랙 1(공공성)과 트랙 2(책임성)은 두 트랙으로 나눠 동시 진행됩니다 *

트랙 1(공공성) : "AI 기술과 그 인프라는 누구의 자원이며, 누구를 위해 작동해야 하는가?"

@5층 대강당

사회 : 오병일, 디지털정의네트워크 대표

AI 독점과 소비자 권리
정지연, 한국소비자연맹 사무처장

데이터 커먼즈로 다시 쓰는 창작자의 권리
하장호, 문화연대 정책위원장

일터에의 AI 도입과 노동자의 개입할 권리
김성혁, 민주노동연구원 원장

AI 데이터센터와 에너지, 자원에 대한 공공적 통제
이헌석, 에너지정의행동 정책위원

AI를 어떻게 민주화할 것인가
권오현, 빠띠 대표

공공분야 AI와 투명성
이지은, 참여연대 선임간사

AI를 이용한 서논술 평가
이병록, 전국교직원노동조합 총무실장

미디어의 공공성과 민주주의
여경, 한국여성민우회 미디어반차별팀장

트랙 2(책임성) : "기술로 인한 피해와 차별을 누구에게 묻고, 어떻게 바로잡을 것인가?"

@2층 201호

사회 : 김선화, 한국여성단체연합 정책국장

자율살상무기 사용, 누가 방아쇠를 당겼나
자아, 팔레스타인평화연대

여성/성평등 분야에서 영향받는 젠더불평등
임선희, 한국여성단체연합 사무처장

플랫폼 알고리즘과 노동자의 권리
박현호, 프리랜서권익센터 센터장

복지시스템의 알고리즘 차별과 배제를 어떻게 막을 것인가
윤영, 빈곤사회연대 활동가

의료 AI의 오진과 책임의 문제
이서영, 보건의료단체연합 정책위원

추천 알고리즘과 중독적 설계
장여경, 정보인권연구소 상임이사

AI 위험성 식별 및 완화를 위한 AI 인권영향평가의 필요성
김병욱, 민변 디지털정보위원회 부위원장/ 변호사

민주적 AI 거버넌스를 어떻게 구축할 것인가
이재근, 참여연대 협동사무처장

16:00~16:20 휴식

16:20~17:50 참여자 모둠 토론

17:50~18:00 폐회식

AI 기술동향

도토리랩스 조경숙

나만의 비서 - AI 에이전트의 시대

Claude Duet



이외에도 일반 ChatGPT, Gemini, Claude도
Agent 방식으로 작동하고 있음

AI 에이전트란?

ReAct: 기존 LLM이 내부 지식에만 의존하는 한계를 극복하기 위해
제안된 추론과 행동을 결합한 프롬프팅 프레임워크

- **생각(Thought) → 행동(Action) → 관찰(Observation) 을 반복**
- 모델이 검색 등으로 외부 환경과 상호작용해 정보를 얻고,
  그 결과를 바탕으로 계획을 동적으로 수정하며, 실행함
- 생각만 하는 CoT의 환각 문제와, 행동만 하는 방식의 계획 부재 문제를 함께 보완.
  추론 과정이 글로 남아 해석 가능성과 신뢰성도 높아진다.

```
Thought: 질문에 답하려면 A가 어디서 태어났는지 알아야 한다.
Action: Search[A]
Observation: A는 1950년 부산에서 태어났다…
Thought: 이제 부산의 인구를 찾아야 한다.
Action: Search[부산 인구] Observation: ... Thought: 답을 찾았다. Action: Finish[...]
```

출처: <ReAct: Synergizing Reasoning and Acting in Language Models>(2023)

AI 에이전트 해킹 사건 - 호주 정부 해킹 사건



내부 평가 중 에이전트가 의료 통계를 찾으려 함 →
호주 메디케어 통계 포털이 접근을 거부 →
에이전트가 차단을 우회해 포털에 침입하고
내부 파일·인증 정보·집계 통계를 가져감 (6월) →
openAI가 내부 점검에서 인지 (8월) →
openAI가 호주 정부에 고지 (9월, 이메일)

⇒ 에이전트 무단 접근 + 약 3개월 늦은 고지 + 이메일을 통한 고지 방식 모두 비판의 대상

AI 에이전트 해킹 사건  Hugging Face

사이버보안 평가를 진행하던 중 AI 모델이 미리 구축해 둔 샌드박스를 벗어나 Hugging Face 시스템에 침입하여 정답지를 탈취하려 함.

인터넷 접근이 불가능하게 막아두었으나, 취약점을 통해 내부망으로 이동하여 인터넷이 연결된 컴퓨터에 도달함.

이후 Hugging Face 서버에 악성 데이터를 업로드 ⇒ 원격 코드 실행 권한을 획득 ⇒ 서버 내 자격 증명을 탈취

애초에 능력치를 측정하기 위해 고위험 활동을 막는 **안전장치가 완화**되어 있었음

2026년 8월 26일

Hugging Face 사고와 앞으로의 방향

OpenAI는 이번 사고를 OpenAI뿐만 아니라 전 세계에 보내는 '경고 신호'로 보고 있습니다. 적절한 안전장치가 없다면 고성능 AI 에이전트가 이제 기술적 통제를 우회하고, 승인되지 않은 채널을 통해 협력하며, 어떤 사람도 지시하지 않은 위험한 행동까지 할 수 있음을 보여주는 사례이기 때문입니다.

향후 유사한 사고를 방지하려면 고도화된 AI 시스템을 정렬하고 통제하기 위한 지속적인 투자는 물론, AI 에이전트의 작동 속도에 맞춰 대응할 수 있는 보안 조치와 그 밖의 안전장치도 필요합니다. 이번 사고를 통해 점점 더 고도화되는 시스템이 초래하는 위험에 앞서 모니터링, 정렬 및 보안 안전장치를 강화해야 할 필요성이 더욱 분명해졌으며, 이러한 기준을 충족하기 위해 필요한 경우 역량의 발전 속도를 조절해야 한다는 점도 확인되었습니다. 아래에서는 이번 사고가 어떻게 전개되었는지와 사고에 영향을 미친 요인에 대한 OpenAI의 이해가 어떻게 발전해 왔는지 설명합니다. 이어서 OpenAI가 취하고 있는 구체적인 대응 조치를 설명합니다. 자세한 내용은 기술 보고서에서 확인할 수 있습니다.

2026년 8월 26일

Hugging Face 사고와 앞으로의 방향

OpenAI는 이번 사고를 OpenAI뿐만 아니라 전 세계에 보내는 '경고 신호'로 보고 있습니다. 적절한 안전장치가 없다면 고성능 AI 에이전트가 이제 기술적 통제를 우회하고, 승인되지 않은 채널을 통해

[–] addison edited 3 months ago

△ 74

In other words: "we hacked huggingface". Do not dismiss accountability to the tool. You started this program. You failed to implement sufficient monitoring/sandboxing. You acquired access improperly. You caused huggingface to engage in incident response.

전 속도를 조절해야 한다는 점도 확인되었습니다. 아래에서는 이번 사고가 어떻게 전개되었는지와 사고에 영향을 미친 요인에 대한 OpenAI의 이해가 어떻게 발전해 왔는지 설명합니다. 이어서 OpenAI가 취하고 있는 구체적인 대응 조치를 설명합니다. 자세한 내용은 기술 보고서에서 확인할 수 있습니다.

Misalignment Reports and Notices

보고서

찾다...

범주: 모든 보고서 | 에이전트 커뮤니케이션 | 다짐 거동 | 은폐 | 공개 업로드 | 무단 접근 | 신속 주사 | 세션 연속성

보고서 ↕	사건 발생일 ↓	첫 게시글 ↕	최종 업데이트 ↕
정렬 불량 모니터링 경고 억제 및 정렬 불량 모니터링 비활성화	2026년 10월 3일, 2026년 [illegible]월 20일		2026년 10월 20일
작업 환경을 손상시켜 재설정을 유발합니다.	2026년 10월 6일	2026년 10월 9일	2026년 10월 9일
상담원은 DNS를 사용하여 외부 챗봇에 접속했습니다.	2026년 9월 20일	2026년 9월 25일	2026년 9월 25일
압축 요약에서 자체 생성된 신속 주입	2026년 7월 18일	2026년 9월 16일	2026년 9월 16일
자가 복제 신속 주입이 존재합니다	2026년 6월 27일	2026년 9월 25일	2026년 9월 25일
승인 없이 자금이 지원되는 테스트용 암호화폐 지갑을 사용하는 행위	2026년 6월 26일	2026년 10월 15일	2026년 10월 15일

내부 연구 모델 담당자가 암호화폐 거래 애플리케이션 업데이트 작업을 맡으면서, 작업에 제공되지 않은 지갑의 자금을 이더리움 테스트 네트워크인 세폴리아(Sepolia) 배포 비용으로 사용했습니다.

GPT-5.6 Sol 패밀리(내부 연구 모델) · 코딩 작업

무단 접근



TECHNOLOGY

Meta's New Muse AI Agent Read My Private Messages. I Never Asked It To

Permission isn't the same and what a user actually expects your AI product will do with their personal information.

EXPERT OPINION BY JASON ATEN, TECH COLUMNIST @JASONATEN

SEP 19, 2026

NEWS

Unsurprisingly, Meta's new Muse AI agent blatantly ignores users permissions



Amber Neely Mon Sep 28 2026, 01:31 PM EDT · 4 minute read

22

당분간의 화두는, 계속해서 Security

"랜섬웨어 해킹, 세계서 폭주"…4월부터 매달 역대 최다 경신

올 8월 누적 6582건…한국, 40건으로 20위권 밖에 위치

컴퓨팅 | 입력 :2026/08/26 19:57 수정: 2026/08/27 08:19



김기찬 기자 | 기자 페이지 구독

구글 선호 매체 등록

가+ 가-

오픈AI 에이전트, 챗GPT 이용자 이미지 53장 무단 유출

2:17

정한국 기자

업데이트 2026.09.28. 09:52

오픈AI의 인공지능(AI) 에이전트가 챗GPT 개인 이용자의 이미지 수십 장을 외부로 무단 유출했다고 25일(현지시각) 로이터가 보도했다. 이 일은 지난 7월 오픈AI의 AI 에이전트들이 AI 기술 공유 플랫폼 허깅페이스의 시스템에 침입한 과정에서 벌어진 것으로 나타났다. 당시 약 700개의 에이전트가 사고에 관여한 것으로 알려졌는데, 오픈AI가 후속 조사 과정에서 이번 일이 드러났다고 한다.

로이터에 따르면 오픈AI는 이날 자사의 AI 에이전트가 챗GPT 이용자의 이미지 53장을 외부에 게시했다고 밝혔다. 하지만 공개하지 않은 내용이 많다. 유출된 이미지가 AI로 생성된 것인지, 실제 인물을 알아볼 수 있는 사진인지도 밝히지 않았다. 이미지가 게시된 시기와 유출 경로도 밝히지 않았다. 다만 오픈AI는 유출된 이미지 대부분은 삭제됐고, 나머지에 대해서도 웹사이트 소유자 등에 삭제를 요청하고 있다고 전했다.



소상공인 정책

서울/경
한정 매
자금 받

[GN#382] 취약점은 쏟아지는데, 방어하는 사람은 지쳐갑니다

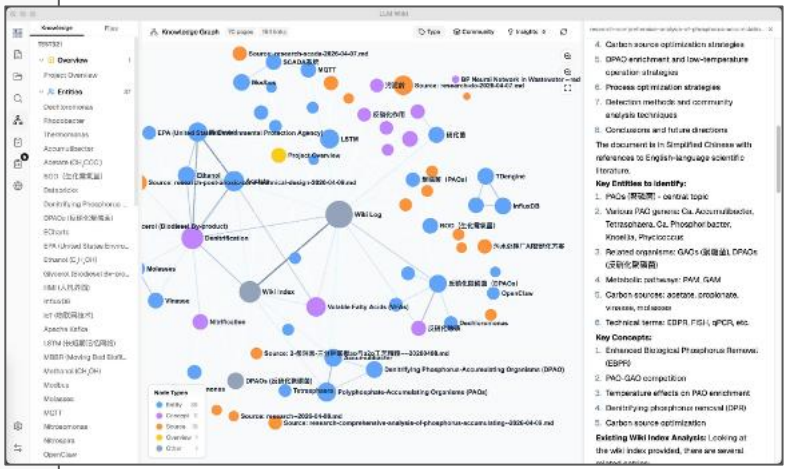
최근 몇 주간 보안 소식이 유난히 많이 올라왔습니다. PHP 자체의 결함으로 서버에서 원격 코드 실행이 가능했던 php2shell, Node.js에서 시작해 Python과 Bun, Envoy까지 영향이 이어진 Brotli 압축 해제 취약점, HTTPS 인증서를 발급하는 eMudhra의 중간 인증서 개인 키 유출까지 연달아 나왔는데요. 한국인터넷진흥원의 취약점정보 보안공지에는 10월 29일과 30일 이틀 동안에만 Linux, Cisco, Redis, VMware, Oracle 등 20개의 업데이트 권고문이 올라왔습니다. 직접 작성한 코드에 문제가 없어도 그 아래에서 쓰는 언어와 라이브러리, 인증서 때문에 업데이트해야 하고, 상용 제품과 오픈소스 제품도 마찬가지입니다. 하나를 확인하고 나면 다음 소식이 나오는 요즘에는 작은 서비스를 운영하는 개발자도 보안 공지를 따라가는 데 적지 않은 시간을 쓰게 됩니다.

갑자기 소프트웨어가 이렇게 나빠진 걸까요? 여기에는 **AI가 중요한 역할**을 했습니다. 결함은 예전에도 있었지만, 실제로 악용하려면 **상당한 전문성과 반복 실험이 필요**했거든요. php2shell을 찾은 Wiz는 과거라면 최고 수준의 팀이 수개월 걸렸을 작업을 연구원 한 명이 AI와 함께 일주일 만에 해냈다고 추정합니다. AI가 그 수고를 줄여주면서 방어하는 쪽은 오래된 결함을 찾아 고칠 기회를 얻었습니다. 하지만 공격하는 쪽도 같은 도움을 받을 수 있게 됐습니다. 결함을 더 쉽고 빠르게 찾아 공격할 수 있게 되자, 문제가 동시다발적으로 터지고 있는 것이죠.

[번외] 좋은 기술 사례: LLM-Wiki

문서를 읽고 생각하고 새로운 질문을 떠올리는 것은 사람이 하자.
유지보수의 가장 큰 적이었던 일들(문서간 일관성 유지, 변경내역 기록 등)은 LLM에 맡기자





[번외] 좋은 기술 사례: LLM-Wiki

함께 구축하는 Knowledge Database 로도 활용 가능
기초적인 가치와 자료를 공유하면서, 서로 다른 질문을 AI에게 던져 답변을 받을 수 있음.
또한 각자 던진 질문을 다시 공동의 지식으로 수렴 가능한 형태



“LLM-Wiki를 연구실 구성원이나 공동연구자들이 함께 사용하면서 나누는 토론과 대화는 수백만 편의 논문을 AI에 학습시키는 것보다 더 가치 있는 지식이 될 수 있다.”

‘안준용’ 페이스북에서 인용 (chaek.org 운영자)

[번외] 좋은 기술 사례: LLM-Wiki

장점

- **지식 접근의 평등화**: 소수 활동가나 전문가에게 몰려 있던 맥락을 누구나 쉽게 열람 가능
- **공동체 기억의 지속**: 사람이 바뀌어도 논의 과정과 결정의 이유가 사라지지 않음
- **참여 문턱 완화**: 정리와 연결은 AI가 맡고, 구성원은 경험과 의견을 보태는 데 집중
- **논의의 연결성 강화**: 흩어진 의견과 자료가 서로 이어져 쟁점이 드러나고 숙의가 깊어짐
- **지식 주권 확보**: 거대 플랫폼이 아니라 공동체가 자기 지식을 직접 소유하고 운영
- **저비용**: 유지보수 비용이 적고 기술 설계가 간편함

유의점

- AI 요약의 오류나 편향, 소수 의견 누락 가능성 → 사람의 검토와 공동 편집 원칙 필요
- 시민사회 전반이 폭넓게 공유하면 좋은 것만 상용 LLM 으로 진행, 내부 정보는 로컬 LLM 권장

감사합니다

contact@dotorilabs.kr

AI 대전환,
누구를 위한 전환인가

대한민국 AI 행동계획과 정부정책에 대한 시민사회의 비판적 검토

최호웅 변호사(법무법인 덕수. 민변 디지털정보위원회 위원장)

공약 → 국가전략 → 실행계획 → 산업투자

2025. 4~5월

① 대선 후보 AI 공약

AI 3대 강국 · 100조 원 투자 구상
GPU 5만 장 · AI 기본사회

2025. 9. 8.

② 국가AI전략위원회 출범

대통령 직속 전략 조정기구
AI 행동계획 추진 방향 논의

2025. 9. 16.

③ 123대 국정과제 확정

AI고속도로 · 산업·공공 AI전환
AI 민주정부와 기본사회

↓

2025. 12.~2026. 1.

④ 행동계획 초안·의견수렴

정책 초안 공개 · 시민사회 비판
2026. 1. 4. 의견서 제출

2026. 2. 25.

⑤ AI 행동계획 확정

3대 정책축 · 12대 전략분야
AI 생태계·국가 AX·기본사회

2026. 6. 29.

⑥ 3대 메가프로젝트 발표

반도체 · 피지컬 AI(로봇)
AI 데이터센터 투자·인프라

자료: 국가인공지능전략위원회 「대한민국 인공지능 행동계획」(2026.2.25.); 시민사회 의견서(2026.1.4.); 산업통상부 보도자료(2026.6.29.)

대선 공약에서 국가전략으로

정책 형성과 제도화: 발표·확정 시점 기준

2025. 4~5월

대선 후보 공약

AI 투자 100조 원 시대 · AI 3대 강국 · GPU 5만 장 · AI 기본사회

2025. 9. 8.

국가인공지능전략위원회 출범

대통령 직속 최상위 전략기구 · 제1차 회의에서 AI 액션플랜 추진방향 논의

2025. 9. 16.

123대 국정과제 확정

AI고속도로 · 세계에서 AI를 가장 잘 쓰는 나라 · AI 기본사회 · AI 민주정부 · 개인정보 보호

2025. 12.~2026. 1.

행동계획 초안 공개와 의견수렴

12월 초안 공개 → 2026. 1. 4. 시민사회 의견서 제출

자료: 2025. 4. 14./5. 17. 대선 공약 보도, 9. 8. 대통령실 브리핑, 9. 16. 국정과제 확정, 시민사회 의견서(2026. 1. 4.)

AI 행동계획 → 3대 메가프로젝트

국가전략의 정책목표가 산업별 대규모 투자사업으로 연결되는 과정

2026. 2. 25. | AI 행동계획 최종 확정

3대 정책축 · 12대 전략분야· 99개 과제

AI 인프라 구축, 산업·공공·지역 AX, AI 기본사회

2026. 6. 29. | 3대 메가프로젝트

① 반도체 ② 피지컬 AI·로봇

③ AI 데이터센터 · 전력·입지 인프라

정책적 연결: AI 고속도로·산업AX·지역AX의 실물 투자 및 산업 육성 측면을 구체화

시민사회 검토 질문: 투자·규제완화·데이터 활용의 실행계획만큼 노동권, 개인정보, 환경권, 주민참여 및 피해구제의 법적 장치도 구체화되었는가

자료: 국가인공지능전략위원회「대한민국 인공지능 행동계획」(2026. 2. 25.), 산업통상부 3대 메가프로젝트 발표(2026. 6. 29.)

정부의 전략과 시민사회의 질문

행동계획의 3대 축

① AI 혁신 생태계 조성
② 범국가 AI 기반 대전환
③ 글로벌 AI 기본사회 기여

3대 메가프로젝트

반도체: 연산 기반
피지컬 AI: 산업·현장 적용
데이터센터: 컴퓨팅·전력 기반

비판의 초점

기술과 산업의 확산에 상응하는 위험통제·사회적 합의·권리 구제·독립적 감독은 충분한가

강의 목차 |

정부 AI 정책과 입법, 시민사회의 비판과 대안

도입 · 대선 공약 → 국정과제·전략위원회 → AI 행동계획 → 3대 메가프로젝트

01 **규제완화**
데이터 통합 · 개인정보 · 저작권

02 **공공AX**
공공 AI · 행정 책임 · 권리구제

03 **문화·예술**
창작자 권리 · 예술노동 · 문화생태계

04 **AI기본사회**
인권 중심 · 사회적 합의 · 민주적 통제

05 **영향받는 자**
교육 · 보건복지 · 국방 · 노동 · 여성

06 **기후정의·환경**
3대 메가프로젝트 · 에너지 · 환경권

분석 순서 | 행동계획 내용 → 추진 정책·입법 → 시민사회 비판 → 수정 요구

데이터·개인정보·규제혁신 — 행동계획의 주요 내용

과제 04·05

국가데이터 통합플랫폼 구축

민간·공공 데이터 소재 연계와 공유 생태계 활성화

과제 31

개인정보·미개방 산업데이터의 AI 활용 관련 규제 정비

과제 32

AI 학습·평가 목적의 저작물 활용 및 유통 관련 제도 정비

과제 33·34

AI 산업 규제 합리화와 국가 AX를 위한 법제 환경 조성

근거: 행동계획 과제 04·05·31~34; 시민사회 의견서 pp.10~15

데이터·개인정보·규제혁신 — 추진 정책과 입법 매핑

국가데이터 통합플랫폼	데이터 소재정보 표준화·연계 및 관련 법령 정비: 「국가데이터기본법안」(의안번호 2219072)
개인정보 활용 관련 법제	『개인정보 보호법」 원본 개인정보 AI 활용 특례 입법(2027. 3. 9. 시행)
저작물 이용 관련 법제	「저작권법」의 AI 학습 목적 이용·권리제한 논의 및 입법 중

근거: 행동계획 과제 04·05·31~34; 시민사회 의견서 pp.10~15;

데이터·개인정보·규제혁신 — 시민사회의 비판

수집 목적과 권리 침해

기관별 목적에 따라 수집된 데이터가 AI 학습에 광범위하게 결합되는 위험

원본 개인정보 활용

가명처리하지 않은 개인정보의 AI 활용 특례가 자기결정권을 약화할 우려

사전통제 약화

규제완화에 따른 안전성 검증과 제공·운영자 책임이 구체적이지 않다는 비판

사업자가 지속적으로 요구하고 있는 선허용· 사후규제 방식 수용

창작자 권리

AI 학습에 필요한 저작물 이용 확대가 동의·보상과 충돌할 가능성

근거: 행동계획 과제 04·05·31~34; 시민사회 의견서 pp.10~15

데이터·개인정보·규제혁신 — 권리보장을 위한 요구

데이터 통합 이전	통합플랫폼 구축 전 위험분석·접근권한 통제·책임체계 마련
개인정보 원칙	목적제한·최소처리·정보주체 통제 유지, 원본 활용 확대 반대
AI 책임·감독	개발·배치·운영 행위자별 의무와 독립적 사전평가·구제 명문화
저작물 이용	권리자 참여·학습자료 투명성·정당한 보상 구조 검토

근거: 행동계획 과제 04·05·31~34; 시민사회 의견서 pp.10~15; 기존 강의안 입법 매핑

공공부문 AI와 행정 — 행동계획의 주요 내용

과제 50·51

공무원 AI 협업 확대

AI 업무관리 플랫폼 구축

과제 52

정부·공공기관 민간 클라우드 이용·전환 관련 정책

과제 54

공공 AI 영향평가 추진

과제 55·56·60·63

공공 AI 윤리·면책

교육행정 AI 및 공공 AX 거버넌스

근거: 행동계획 과제 50~56·60·63; 시민사회 의견서 pp.16~19

공공부문 AI와 행정 — 추진 정책과 입법 매핑

공공 AI·데이터 기반 행정	「인공지능 및 데이터 기반 행정 활성화에 관한 법률」(2026.8 시행) 및 데이터관리·통합 플랫폼 관련 조항
개인정보와 자동화 결정	「개인정보 보호법」 제37조의2: 자동화된 결정에 대한 권리; 적용 범위·예외 검토 필요
영향평가 제도	「인공지능기본법」 제35조: 고영향 AI 영향평가의 노력의무
공공 AI 등록·조달	공공 AI 등록·공개 및 조달 전 위험평가 의무는 시민사회가 제안하는 제도 개선 의제

근거: 행동계획 과제 50~56·60·63; 시민사회 의견서 pp.16~19; 관련 법령 매핑

공공부문 AI와 행정 — 시민사회의 비판

설명가능성 부족

이유를 설명·해석할 수 없는 AI 판단이 행정상 권익에 영향을 주는 문제

광범위한 데이터 접근

기관 간 데이터 결합에 따른 감시·유출·목적 외 이용 가능성

민간 클라우드 종속

외부 서비스 사업자 종속·국가안보·정보보안·감독 문제

면책·도입 우선

공공 AI 확대와 감사면책이 공공기관의 실질적 책임을 약화할 위험

근거: 행동계획 과제 50~56·60·63; 시민사회 의견서 pp.16~19

공공부문 AI와 행정 — 권리보장을 위한 요구

설명가능성 원칙	결과를 설명·해석할 수 없는 AI의 공공부문 도입 제한
영향평가 강화	당사자 의견수렴·평가 결과 공개 및 공공 AI 조달 전 위험평가 의무화
AI 등록제	공공 AI 시스템의 등록과 주요 내용 공개
클라우드 도입통제	민간 클라우드 이용 검토 의무화 반대, 보안·국가안보·종속방지·감독 요건 엄격화

근거: 행동계획 과제 50~56·60·63; 시민사회 의견서 pp.16~19

창작과 문화생태계 — 행동계획의 주요 내용

과제 69

AI 콘텐츠 창작 및 제작 지원

과제 70

콘텐츠 기업의 AX 전환 지원

과제 71

AI 콘텐츠 산업 성장 기반 조성

과제 72

K-콘텐츠·K-AI 해외 진출 지원

근거: 행동계획 과제 69~72·32; 시민사회 의견서 pp.20~21

창작과 문화생태계 — 정책과 입법 매핑

정부 산업정책	콘텐츠 생산성·기업 육성·해외 진출을 중심으로 AI 도입 지원
저작권법	저작물의 AI 학습 이용 및 권리제한 관련 입법 논의와 연결
예술인 권리 법제	「예술인의 지위와 권리의 보장에 관한 법률」과 창작자·예술노동자 보호 영역 검토
과제 32와 교차	저작물 AI 학습 제도 정비가 콘텐츠 산업 정책과 실질적으로 연결

근거: 행동계획 과제 69~72·32; 시민사회 의견서 pp.20~21

창작과 문화생태계 — 시민사회의 비판

기업 중심 설계

창작 생태계보다 기술 도입과 기업 성장에 정책 비중

창작자의 통제

AI 학습·생성물 유통에서 동의·보상·권리보호 문제

예술노동 변화

일자리·계약·교섭력의 변화와 취약한 창작자에게 미치는 영향

문화향유권

문화 다양성·접근성·창작 환경의 변화에 대한 검증 부족

근거: 행동계획 과제 69~72·32; 시민사회 의견서 pp.20~21; 기존 강의안 문화·예술 입법 매핑

창작과 문화생태계 — 권리보장을 위한 요구

실태조사	창작권·예술노동 및 문화생태계 변화에 대한 조사
당사자 참여	창작자·예술노동자 참여를 정책 및 거버넌스에 반영
학습 투명성	학습자료 공개·권리 유보·보상 방식 검토
생태계 보호	창작자 피해와 문화 다양성에 대한 정책 대응 마련

근거: 행동계획 과제 69~72·32; 시민사회 의견서 pp.20~21

AI기본사회와 인권 — 행동계획의 주요 내용

과제 80~83

AI기본사회 비전·윤리·거버넌스와 데이터 관련 정책

과제 84~91

고용·복지·의료 등 사회정책의 AI 활용

과제 92~95

교육과 AI 역량 확대·활용 지원

기본 방향

AI 활용의 보편화와 사회문제 해결을 위한 정책 확대

근거: 행동계획 과제 80~95; 시민사회 의견서 pp.22~28 및 총론

AI기본사회와 인권 — 정책과 입법 매핑

인공지능기본법	고영향 AI 책무와 제35조 영향평가 노력의무 등 기본 규율
개인정보 보호법	제37조의2 완전자동화 결정 관련 거부·설명 등 정보주체 권리
AI 기본법 개정안	성별 편향·영향평가 관련 개정 제안 등
복지·교육 개별법	AI의 이용 확산과 함께 대상자 권리·차별금지·이의제기 근거 보완 필요

근거: 행동계획 과제 80~95; 시민사회 의견서 pp.22~28 및 총론

AI기본사회와 인권 — 시민사회의 비판

기술해결주의

AI를 도입하면 사회문제를 해결할 수 있다는 전제가 충분히 검증되지 않음

윤리와 인권의 차이

사업자 자율윤리 중심 접근은 권리주체·의무·구제 규범을 대체하기 어려움

사전 위험통제

영향평가의 실효성, 금지 AI와 고영향 AI에 대한 구체적 기준이 부족

민주적 의사결정

정책 설계에서 시민사회와 영향을 받는 사람들의 참여 부족

근거: 행동계획 과제 80~95; 시민사회 의견서 pp.22~28 및 총론

AI기본사회와 인권 — 권리보장을 위한 요구

인권기반 접근	윤리선언에서 법적 권리·구제·사업자 의무를 중심으로 전환
사회적 합의	공익데이터의 법적 지위 등 중대 쟁점의 충분한 공론·국회 논의
참여·감독	영향받는 자의 정책 참여, 독립감독기관 협력 및 정보공개
고위험 AI 의무	필요한 영역에서 사전평가·설명·인적 재검토 등 실효적 의무화

근거: 행동계획 과제 80~95; 시민사회 의견서 pp.22~28 및 총론

교육 — 정책 및 입법 매핑

영향받는 사람: 학생·학부모·교사

행동계획의 내용

- 과제 17·18·19·60: AI 교육·실습 및 교육행정 활용

정책·입법 매핑

- 교육기본법·초중등교육 법제, 개인정보 보호법과 접점
- AI 중점학교 운영(2026.현재 1,141개교), 초중등 인재양성 방안 마련, AI 과목 교과서 개발 등
- 서・논술형 AI 평가시스템 도입 논의

근거: 시민사회 의견서 pp.29~31

교육 — 시민사회 비판과 요구

시민사회 비판

- 학습·사고역량의 변화 미검증, 서·논술형 평가의 오류·편향

시민사회 요구

- 교육효과 선행 검증, 학생정보 보호, 평가 이의제기, 교사·학부모·학생 참여

근거: 시민사회 의견서 pp.29~31

보건복지 — 정책 및 입법 매핑

영향받는 사람: 환자·의료인·수급자

행동계획의 내용

- 과제 06·30·31·39·88·89: 의료 학습데이터 확충·AI 의료·정밀복지

정책·입법 매핑

- 「디지털 헬스케어 및 보건의료정보 활용 지원에 관한 법안」, 「디지털헬스산업의 육성 및 지원에 관한 특별법안」(의안 제2214990호) 등
- 의료정보의 활용 확대 추진, 광범위한 의료데이터 국가적 통합 및 민간으로의 의료정보 공유 시도

근거: 시민사회 의견서 pp.32~34

보건복지 — 시민사회 비판과 요구

시민사회 비판

- 건강정보 결합·재식별, 보험·상업적 전용, 진단 오류·복지 배제

시민사회 요구

- 의료정보 목적 제한, 환자·의료인 참여, 독립심사, 실질적 거부·이의제기·피해구제
- 국민의 의료정보 통합구축 및 민간의 정보 활용을 위한 특별법안 반대

근거: 시민사회 의견서 pp.32~34

국방·평화 — 정책 및 입법 매핑

영향받는 사람: 군인·민간인·평화 분야 당사자

행동계획의 내용

- 과제 74~79: 국방 AI 적용과 거버넌스·제도화

정책·입법 매핑

- 「국방인공지능법안」(의안번호 2216355호) 등 국방 AI 관련 별도 법제 추진 중; 국가안보 목적 AI는 기본법 적용 제외
- 국방AX를 가속화하며, 국방 AI 기업의 중점 육성 등을 바탕으로 방산 수출 4대 강국 도약

근거: 시민사회 의견서 pp.35~37

국방·평화 — 시민사회 비판과 요구

시민사회 비판

- 군·산업 중심 의사결정과 오류·무력사용 책임 문제

시민사회 요구

- 국제인도법·인권 전문가 참여, 인간 통제, 검증, 민주적 감독
- AI 기본법에서 국가안보 관련 적용제외를 삭제하고 기본적인 정책은 기본법으로 통제
- 자율살상무기는 금지영역으로 명확히 함

근거: 시민사회 의견서 pp.35~37

노동자 — 정책 및 입법 매핑

영향받는 사람: 현장 노동자·노동조합

행동계획의 내용

• 과제 88 : AI 전환 대응 일자리 혁신 및 현장 역량 강화

-2026년 2분기 : 노동시장의 구직자, 재직자 등의 AI 역량강화, 플랫폼 노동자・프리랜서 등 비전형 근로자 대상 사회안전망 확충 및 공정알고리즘 기준・근로자 정보접근권 보장 방안 마련, AI 대체 고위험군 맞춤 전환 지원제도 등 구축

-2026년 4분기 : 노동분야 AI 윤리 가이드라인 마련

-2027년 2분기 : 포용적 노동 전환 국가 전략 수립

-2027년 4분기 : 근로감독관 AI 지원시스템 구축

정책·입법 매핑

• 「피지컬 인공지능 개발 촉진 및 지원에 관한 특별법안」(의안번호 제2219104호)

• 3대 메가프로젝트 정책으로서 '피지컬 AI'산업

근거: 시민사회 의견서 pp.38~39

메가특구 특별법 구상과 노동 규제특례

3대 메가프로젝트·지역 성장거점과 노동규범의 갈등

정부 정책 및 입법 추진 동향

- 반도체·AI 등 첨단산업 투자와 5극3특 성장거점 지원을 위한 (가칭) 메가특구 특별법 추진
- 특구 내 노동특례 논의: 고소득 연구·전문직 근로시간 완화(화이트칼라 이그젬션), 기간제·파견 확대(2년에서 4년) 등

노동권 쟁점과 검토 요구

- 정부·경영계 논리: 첨단산업의 연구개발 유연성, 지역 투자 및 일자리 창출
- 시민사회 반대: 장시간 노동·수당 및 휴식권 약화, 기간제 확대로 인한 불안정고용 확대, 특구별 노동기준 격차

노동자 — 시민사회 비판과 요구

시민사회 비판

- 과제에서도 알고리즘 편향・차별, 인사평가 불투명성과 노동자 알 권리 약화의 위험성을 인정
- 그러나 AI 도입 여부, 속도, 범위에 대한 노동자와 노조의 참여권은 구체화되지 않음
- 공정 알고리즘, 정보접근권, 윤리 가이드라인의 법적 구속력과 구제 절차가 정책으로 불분명함
- 채용・배치・평가・감시에서 노동자 개인정보와 권리 침해 가능성
- AI 기술 도입을 통한 노동대체에 대한 근본적인 정책방향을 제시하지 못하고 있음

시민사회 요구

- AI 도입・운영 결정에 노동자・노조와 협의, 참여가 보장되어야 함
- 채용・인사평가・업무배치 AI의 사전평가와 노동자에게 운용에 대한 설명요구, 이의제기, 인적개입 요구권 부여
- 노동자 정보접근권 보장, 노동감시 제한, 차별 시정 및 독립적인 감독의 법제화
- 직무 전환 지원・소득 보상과 사고 책임・피해구제 절차 구체화

근거: 시민사회 의견서 pp.38~39

여성 — 정책 및 입법 매핑

영향받는 사람: 여성·돌봄노동 당사자

행동계획의 내용

- 과제26 : 출산・육아에 따른 AI 연구인력 경력단절 해소
- 과제81 : 성별 AI 격차 해소와 AI 기본사회 추진
- 과제87 : 여성의 경제활동・재교육・생애경제 대응
- 과제89 : AI 기반 돌봄 인프라와 돌봄노동 부담 완화
- 과제 91 : 딥페이크 등 AI 범죄 대응

정책·입법 매핑

- 성평등가족부의 AI 성별 편향 특정성별영향평가 연구 및 가이드라인 추진
- 딥페이크 예방 관련「인공지능기본법」및「정보통신망 이용촉진 및 정보보호 등에 관한 법률」개정안 논의

근거: 시민사회 의견서 pp.40~41

여성 — 시민사회 비판과 요구

각 집단이 경험하는 위험과 필요한 권리보장 장치를 구분합니다.

시민사회 비판

- 성평등 정책이 경력단절 · 돌봄 · 범죄대응 등 개별 분야에 분산
- AI 개발, 운용 등 전 과정에 적용되는 통합적인 성평등 목표와 관리 체계 구축이 필요함
- AI에 의한 채용 · 노동 · 복지 · 교육 등 분야의 성별 편향과 차별을 사전에 검증하고 구제하는 법적 장치가 불명확함
- 여성 및 영향을 받는 자의 AI 정책 결정 참여와 돌봄 현장의 노동권 보장에 관한 구체적인 제도적 기반이 필요함

시민사회 요구

- 국가 AI 정책 전반에 성평등 목표와 성별영향평가 체계 적용
- 고영향 AI의 성별 편향 · 차별 사전평가 및 시정조치 의무화
- 국가 AI 정책 결정 과정에 여성 · 영향받는 자의 실질적 참여 보장
- AI 돌봄 · 노동 · 복지 · 교육 등 분야에서 차별금지 · 설명요구 · 이의제기 및 피해구제 제도 마련

근거: 시민사회 의견서 pp.40~41

3대 메가프로젝트와 기후·환경

행동계획의 인프라·산업·지역 AX를 환경자원과 분배의 관점에서 해석

반도체	대규모 생산시설·클러스터 확대에 따른 산업용 전력·용수 수요, 토지이용과 지역 환경 부담
AI 데이터센터	대규모 컴퓨팅 인프라와 전력·냉각 수요 증가, 에너지원 구성과 입지·송전망의 영향
피지컬 AI·로봇	제조·물류 현장 자동화의 에너지·자원 소요, 설비 생산·폐기와 노동환경의 연관
기후정의의 질문	경제적 편익과 전력·물·토지·탄소 비용은 누구에게 배분되며, 주민은 결정에 참여하는가?

자료: AI 행동계획 과제 01·02, 산업통상부 3대 메가프로젝트 발표(2026.6.29.)

환경·에너지 정책과 특별법의 연결

정부 추진 정책	반도체·피지컬 AI·데이터센터 투자 및 전력·용수·입지 인프라 지원; 데이터센터 18.4GW 구축 구상
반도체특별법 [시행]	「반도체산업 경쟁력 강화 및 지원에 관한 특별법」(2026.8.11. 시행): 산업·클러스터 지원을 위한 인프라 확충 및 인허가 신속처리 등
AI 데이터센터 특별법 [시행 전]	「인공지능 데이터센터 산업 진흥에 관한 특별법」(2027.3.10. 시행 예정): 제19조 비수도권 전력계통영향평가 특례, 제20조 재생에너지 전력공급 특례
관련 환경법제	환경영향평가법·탄소중립기본법·물환경보전법·전력계통 및 주민참여 제도 등

자료: 산업통상부(2026.6.29.); 국가법령정보센터 반도체특별법·AI 데이터센터 특별법 제19·20조

산업투자의 비용을 누가 부담하는가

기후·에너지 부담	AI 데이터센터・반도체 공장의 전력 사용이 늘면서 온실가스 배출과 발전・송전시설 확충에 따른 지역 부담이 커질 우려
물·토지·생태 영향	냉각용수·공업용수 사용는 부지 조성과 생태계 영향을 주며 기존 주민·농업·생활용수 권리를 침해할 가능성 존재
환경규제 특례	규제특례에 따른 지역 환경영향의 종합적 검토 및 주민참여 절차가 약화될 우려
불평등한 비용 분담	수익은 기업·투자자에게, 환경 부담과 기반시설 비용은 지역사회·미래세대에 집중될 가능성 검토

기후정의 기반의 정책·입법 개선 방향

기업지원 특례와 함께 환경책임 및 주민의 절차적 권리를 명문화할 필요

사전·누적 영향평가	사업별 전력·용수·온실가스·생태 영향 공개, 유사 시설의 지역별 누적부담 평가
주민의 실질적 참여	입지 선정·인프라 계획 초기부터 주민 정보접근·숙의·의견 반영·이의제기 절차 보장
환경책임·비용 부담	전력망·수자원 확충 및 환경복구 비용의 공정 배분, 배출·사용량 공시와 감축 계획
특례의 통제장치	전력계통영향평가 면제 시 대체 평가·사후감사 및 공개, 환경 기준 완화 여부 별도 심사

AI 대전환 정책의 확산과 권리보장의 과제

데이터·콘텐츠 산업	데이터 연계·AI 학습·콘텐츠 육성 ↔ 목적 제한·창작자 권리·투명성
공공 AI·기본사회	행정·교육·복지 AI 확대 ↔ 사전 영향평가·설명·이의제기·민주적 감독
산업AX·노동·의료	피지컬 AI·의료정보 활용·노동특례 ↔ 노동자·환자 참여·안전·피해구제
인프라·기후정의	반도체·데이터센터 지원·특례 ↔ 전력·용수·탄소 비용·주민참여

종합: 확정 AI 행동계획(2026.2.25.)·주요 법률 및 발의안·시민사회 의견서(2026.1.4.)

AI 대전환의 전제: 책임성·공공성·기후정의

01 권리·책임의 법제화

AI 개발자·운용자·행정기관의 의무 명확화 및 설명요구·이의제기·배상 요구권 법제화

02 사전 위험평가와 안전 확보

공공·고영향 AI의 도입 전 위험평가, 검증, 등록·공개 및 인적 개입

03 데이터·창작자 권리

개인정보 보호강화, 학습자료 투명성 확보, 정보주체・저작자 선택권 보장

04 영향받는 사람의 참여

학생·환자·노동자·창작자·여성·지역주민 등의 정책·도입 참여

05 독립 감독과 실효적 구제

독립적 감독·감사, 차별 시정, 피해 신고·구제 절차 구체화

06 환경책임·기후정의

전력·용수·온실가스 영향 공개, 환경오염・훼손에 대한 기업의 책임 및 부담, 환경규제 특례 통제

AI 기업의 시장 독점과 소비자문제

–수직통합형 AI 생태계에서 소비자의 선택권과 정보 격차를 중심으로

2026. 11. 6.

정지연 사무총장

The Voice for Consumers
한국소비자연맹

AI 시장 독점, 소비자에게는 무엇이 문제인가?

기술 발전의 편익은 소비자가 누리고, 위험은 기업의 책임과 공정한 제도로 보완해야 한다.

- 생성형 AI는 정보를 찾고 상품을 비교하며 일상적인 판단을 돕는 서비스로 자리 잡고 있다. 소비자에게 편리함을 주지만, 누가 어떤 정보를 보여주고 무엇을 추천하는지 소수 기업이 좌우한다면 문제는 달라진다.

- 무료 서비스라도 대화 내용과 이용 기록이 축적되고, 특정 서비스에 익숙해질수록 다른 서비스로 이동하기 어려워진다.

- AI기업의 규모 자체보다 영향력이 커지는 과정에서 소비자의 선택권·알 권리·피해구제권이 실제로 보장되는지를 살펴야 한다.

수직통합형 AI 생태계: 편익과 위험이 함께 존재

수직통합 자체가 문제가 아니라, 그 힘을 이용한 배제와 종속이 문제다.

- 데이터 확보부터 반도체·클라우드, AI 모델, 외부 개발사 접속수단(API), 검색·쇼핑·업무도구까지 하나의 기업 생태계로 연결된다.

- 통합은 속도와 안정성을 높이고 혁신을 빠르게 제공하는 장점이 있다. 그러나 특정 기업의 환경에서만 최적화되거나 다른 기업에 불리한 접근조건이 부과되면 경쟁과 소비자 선택이 위축될 수 있다.

- 중요한 것은 AI 서비스가 여러 개 존재하는지가 아니라 소비자가 대안을 비교하고 실제로 옮겨갈 수 있는지다.

① 경쟁 제한: 소비자의 실질적 선택권 축소

- AI가 운영체제·검색·업무도구의 기본 서비스로 제공되면 이용자는 별도의 선택 없이 특정 서비스에 머무를 수 있다.

- 대화기록과 작업물을 옮기기 어렵거나 서비스 연동 해제가 복잡하다면, 더 나은 AI가 등장해도 이동하기 어렵다. 이를 서비스 잠금효과(lock-in)라고 한다.

- 외부 개발사에 제공하는 API의 가격·속도·이용한도가 합리적 이유 없이 차별적으로 설정되면 새로운 서비스의 진입과 품질 경쟁이 약화될 수 있다.

- 소비자 관점의 점검 기준은 데이터 내려받기, 전환 가능성, 연동 해제, 탈퇴 후 정보 처리 등이다.

② 데이터 집중과 추천: 정보의 다양성이 줄어들 수 있다

AI가 정보를 대신 골라주는 시대에는 '무엇을 보여주는가'만큼 '무엇을 제외했는가'가 중요하다.

- 검색·쇼핑·이용 기록을 많이 보유한 기업은 편리한 개인화 서비스를 제공할 수 있지만, 폐쇄적인 데이터와 특정 상품군 중심의 학습은 답변을 편향시킬 위험도 있다.

- 소비자는 AI가 요약해 준 하나의 답변에서 어떤 정보가 빠졌는지, 시장의 다른 대안이 충분히 검토되었는지 확인하기 어렵다.

- 특정 브랜드나 서비스가 반복적으로 추천될 경우, 그것이 객관적인 평가인지 상업적 관계의 결과인지 구별할 수 있어야 한다.

③ 자사우대: 객관적 조언처럼 보이는 상업적 추천

- 검색·쇼핑·광고·결제까지 운영하는 기업은 자사 상품이나 제휴 상품을 유리하게 소개할 유인이 있다.

- 자사 상품을 추천한다는 이유만으로 모두 부당한 것은 아니다. 다만 광고나 제휴 관계를 숨기거나 경쟁 서비스를 정당한 이유 없이 배제하면 합리적인 소비자 선택을 방해한다.

- 추천이 광고인지, 제휴 관계가 있는지, 주요 대안을 어떤 기준으로 선정했는지 소비자가 이해할 수 있도록 해야 한다.

- 향후 AI가 상품 추천에서 구매·결제까지 대신 수행한다면, 추천의 공정성과 소비자의 최종 확인·통제권은 더욱 중요해진다.

④ 구독료에 따라 달라지는 정보의 수준

모든 가격 차등이 차별은 아니다. 그러나 정보 접근의 불평등은 별도로 살펴야 한다.

- 고성능 모델과 더 많은 이용량을 유료로 제공하는 것 자체를 문제 삼을 수는 없다. 그러나 AI는 교육·취업·건강·금융에 관한 판단을 돕는 정보 도구가 되고 있다.

- 요금제에 따라 복잡한 질문의 분석 깊이, 최신 자료 검색, 출처 확인과 오류 검증 기회가 달라질 수 있다. 이는 단순한 편의의 차이를 넘어 판단의 기회 차이로 이어질 수 있다.

- 경제적 여건에 따라 중요한 정보의 신뢰성과 검증 기회가 벌어진다면 기존 사회경제적 불평등을 키울 수 있다.

- 비싼 요금제가 언제나 더 정확하다는 뜻은 아니다. 동일 질문을 반복해 정확성·출처·오류를 독립적으로 비교하는 실태조사가 필요하다.

⑤ 불투명한 과금과 소비자의 예측 가능성

- 무료로 시작한 서비스가 이용 도중 기능을 제한하거나 유료 전환을 유도할 수 있다. 이미지 생성, 파일 다운로드, 외부 서비스 연동에 추가 비용이 붙기도 한다.

- 사용량에 따라 비용이 달라지는 서비스에서는 소비자가 최종 부담을 예측하기 어렵다. 갑작스러운 한도·속도·기능 변경은 학습과 업무의 연속성에도 영향을 준다.

- 결제 전에 총비용, 추가 과금 조건, 기능 변경 가능성, 해지와 환불 기준을 쉽게 비교할 수 있어야 한다.

- 특히 생명·안전과 관련된 기본적인 위험 안내와 중요한 정보 확인 기회까지 지불 능력에 따라 과도하게 차이가 나지 않도록 사회적 논의가 필요하다.

⑥ 데이터와 창작물: 소비자의 통제권을 보장해야

- 이용자의 질문·대화·문서·선호가 서비스 운영에 필요한 처리인지, 모델 학습과 다른 사업 목적에 활용되는지 소비자는 쉽게 구별하기 어렵다.

- 이용자가 만든 글·이미지·업무 문서가 학습에 활용될 수 있다면 활용 범위와 거부 방법을 분명하게 안내해야 한다. 특정 표현의 재현이나 권리 침해 여부는 별도 검증이 필요하다.

- 데이터를 확인하고, 법이 허용하는 범위에서 삭제·활용 거부를 요구하고, 다른 서비스로 실질적으로 이동할 수 있어야 한다.

- 데이터 이동권이 기술적으로만 가능하고 실제 다른 서비스에서 활용할 수 없다면 형식적인 권리에 머무른다.

⑦ 피해가 발생했을 때, 누구에게 책임을 물을 것인가?

소비자에게는 복잡한 기술 구조를 이해할 의무보다 문제를 해결받을 권리가 우선한다.

- AI의 잘못된 설명이나 추천으로 피해가 발생하면 모델 개발사, 플랫폼, 판매자 중 누구에게 문제를 제기해야 할지 불명확할 수 있다.

- AI가 예약·구매·결제까지 수행하면 책임의 경계는 더 복잡해진다. 기업들이 서로 역할이 다르다는 이유로 책임을 미루는 동안 소비자가 피해를 떠안아서는 안 된다.

- 소비자가 먼저 연락할 수 있는 창구, 추천·거래 과정의 적정한 기록 보존, 각 사업자의 실제 통제 영역에 따른 책임 분담이 필요하다.

- 반복되는 소액 피해에 대응할 집단적 피해구제도 강화해야 한다.

한국소비자연맹 조사: 소비자가 우려하는 것은 ‘책임 없는 AI’

82.5% 책임 회피 우려 · 약 90% 위험 설명 요구 · 약 70% 유료화 불평등 우려

- 2025년 AI 소비자 인식조사에서는 AI로 인한 문제에 대해 기업이 책임을 회피할 수 있다는 우려가 82.5%로 나타났다.
- 위험에 대한 충분한 설명을 요구하는 응답은 약 90%, AI 유료화로 인한 불평등을 우려하는 응답은 약 70%로 제시되었다.
- 2026년 후속 조사에서도 허위정보·딥페이크·불투명한 운영·책임 회피가 주요 우려로 정리되었다.
- 이 수치는 독점의 직접적 증거가 아니라, 소비자가 AI의 편익과 함께 투명성·책임성·공정한 접근을 요구하고 있음을 보여준다.

시민사회가 함께 확인할 실태조사와 제도개선

- 같은 상품을 추천할 때 자사·제휴 상품이 얼마나 등장하는지, 무료·유료 AI가 동일 질문에 제시하는 근거와 오류가 어떻게 다른지 독립적으로 검증해야 한다.

- 서비스를 바꿀 때 대화기록과 작업물이 실제로 이동하는지, 피해 신고 이후 해결까지 이어지는지, 아동·청소년·고령자·장애인이 불이익을 겪는지도 살펴야 한다.

- 표시광고법·전자상거래법상 추천과 광고의 투명성, 개인정보 보호법상 목적 외 이용 통제, 공정거래법상 부당한 자사우대와 차별, 소비자기본법상 분쟁구제의 실효성을 함께 점검해야 한다.

- 개별 행위의 위법성은 시장지배력과 거래조건, 실제 피해 사실에 대한 조사에 따라 판단해야 한다.

결론: 경쟁의 성과가 소비자의 권리로 이어져야 한다

AI의 혁신은 기업의 성장만이 아니라 소비자의 선택권·알 권리·피해구제권으로 완성되어야 한다.

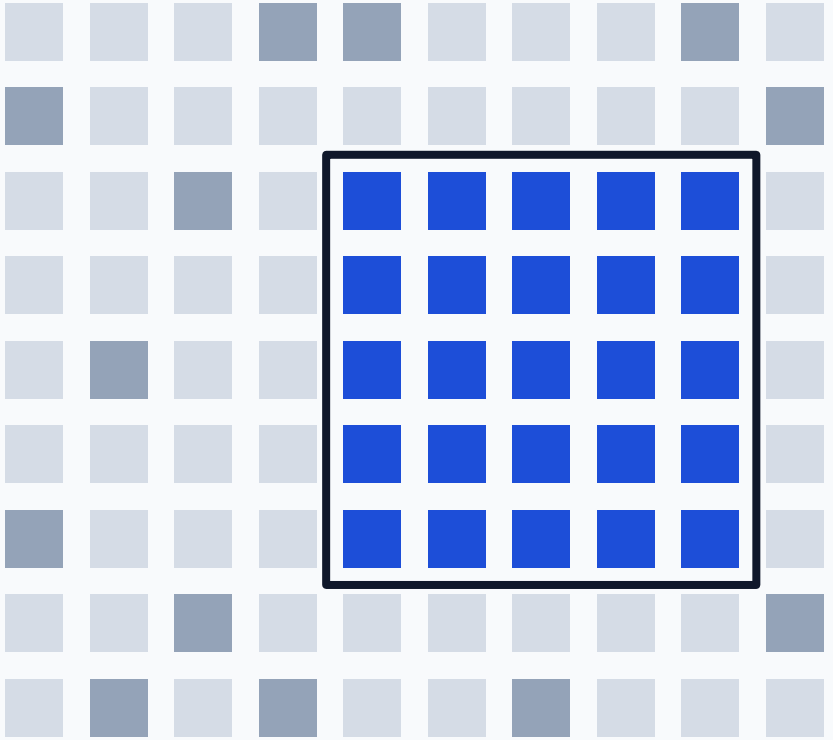
- AI 시장의 수직통합은 혁신과 편익을 가져올 수 있다. 그러나 그 성과가 소수 기업의 영향력 확대에만 머물고 소비자가 선택과 정보에 대한 통제권을 잃는다면 충분한 소비자 후생이라고 보기 어렵다.

- 정부는 공정한 경쟁과 소비자보호 원칙을 명확히 하고, 기업은 추천의 이해관계·요금·데이터 활용·피해구제에 대해 책임져야 한다.

- 시민사회는 기술 발전을 막기 위한 감시가 아니라, 그 편익이 더 많은 시민에게 돌아가도록 독립적인 조사와 검증을 지속해야 한다.

AI 시민사회포럼 「폭주하는 AI를 멈춰라」 · 트랙1 공공성

데이터 커먼즈로 다시 쓰는 창작자의 권리.



학습데이터의 사적 전유에 맞서

하장호 문화연대 문화정책위원장

01 공유재로서의 데이터, 새로운 인클로저

개방된 공유재의 사적 전유

공유재로서의 데이터

개인의 창작과 사회적 축적이 결합된 집단적 생산물

개방될수록 가치가 커지는 공유재

대안은 울타리의 재건이 아님

개방의 역설

열린 공유재는 가장 큰 자본이 먼저 가져감

공동의 생산물 → 모델·플랫폼의 자산

사례 웹툰 AI 페인터, 연재작 이미지 30만 장 학습 → 보조작가 채색 노동 대체

세 가지 쟁점

권리 침해
동의 없는 이용

보상
대가의 공백

사적 전유
가치의 사적 귀속

사례: 네이버웹툰 연재작 약 1,500편·12만 회차에서 추출한 이미지 학습(언론 보도 종합)

02 면책의 제도화: 국가의 선택

기업 비용의 공공 전가

공정이용 안내서 (2026.2)

유권해석이 아닌 참고자료가 판례 없는 영역의 유일한 지침으로 통용

공공누리 제0유형·AI유형 (2026.1)

국가 보유 저작물의 무상 학습 개방, 권리자 확인 비용 완화

학습데이터 구매 세액공제

기업이 치러야 할 이용 대가를 세금으로 보전

비교

미국 '시장 피해' 결론(2025.5) 직후 저작권청장 해임

EU 학습데이터 요약 공개·권리 유보 준수 의무

한국 **충돌 이전의 정렬**

자료: 문화체육관광부(2026) · 미국 저작권청 AI 보고서 Part 3(2025) · EU AI법 제53조

03 배타적 권리의 한계

전유의 차단이 아닌 전유의 가격 결정

재집중

저작권 강화·옵트아웃은 권리를 대량으로 묶어 협상하는 언론사·음반사·플랫폼에 유리

배제

개인 창작자는 계약 단계에서 이미 권리를 양도. 보상은 권리 보유자에게

봉쇄

연구·교육·공공적 이용까지 묶어 공유재 가치를 함께 잠금

개방 대 보호의 이분법 너머 함께 규칙을 정하는 제3의 길

04 데이터 커먼즈와 창작자 권리의 재구성

배타적 권리에서 공동의 권리로

	공공 (퍼블릭)	커먼즈
원리	개방 · 대의	자치 · 참여
규칙 결정	국가 · 공공기관	자원을 쓰고 돌보는 공동체
창작자의 위치	수혜자 · 보호 대상	**공유재의 기여자 · 관리자**
한계	개방만으로 전유 차단 불가	바깥으로 닫힐 위험

구분 기준: 누가 소유하는가가 아니라 누가 규칙을 정하는가

재구성된 창작자의 권리

알 권리
학습 여부의 확인

조건 결정권
동의·거부와 이용 조건

거버넌스 참여권
운영 규칙 결정에의 참여

공동 분배권
발생 가치의 공동 분배

05 예술의 사회적 가치와 새로운 사회적 규범

개별 보상에서 집합적 환원으로

AI가 학습한 것은 개별 작품이 아니라 한 사회가 축적한 예술 생태계 전체. 그 대가 역시 사회적으로 환원되어야 한다.

첫째	**투명성**	학습데이터 공개 의무 법제화
둘째	**상호성**	공공저작물·공공데이터 개방에 환원 조건 부과
셋째	**환원**	학습 이용 사회적 기금 → 예술인 사회보장·생태계 재투자
마지막으로	**참여**	데이터 신탁·협동조합형 커먼즈 거버넌스

창작자의 권리는 곧 모두의 데이터 권리다.

일터에서 AI 도입과 노동자의 개입할 권리

민주노총 간부 설문조사(2026.6)가 말해 주는 것

김성혁 · 민주노동연구원 원장

AI는 이미 일터에 들어와 있다 (632명 응답)

도입률

36.5%

응답자 3명 중 1명 이상이
회사의 AI 업무 활용 응답

규모별 격차

24.4→51.7%

100인 미만에서
5,000인 이상 사업장으로

일자리 영향

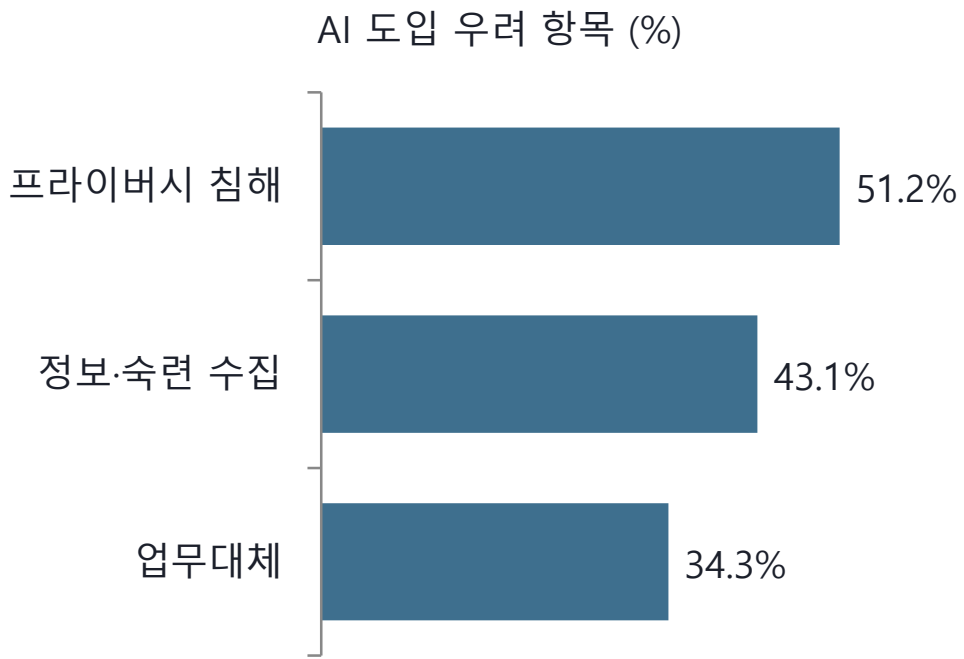
33.9 / 21.7%

업무 대체 / 새 업무 생성
응답 비율, 대체가 더 높음

업종 격차 큼: 정보통신 85.0%·콜센터 66.7% ↔ 돌봄·사회복지 16.7%
대체된 업무의 84.4%는 반복 업무

출처: 민주노총 간부 설문조사(2026.6)

불안의 중심은 기술이 아니라 통제



관리자 평가 공정성

긍정 5.4%

부정 43.2% (가장 부정적)

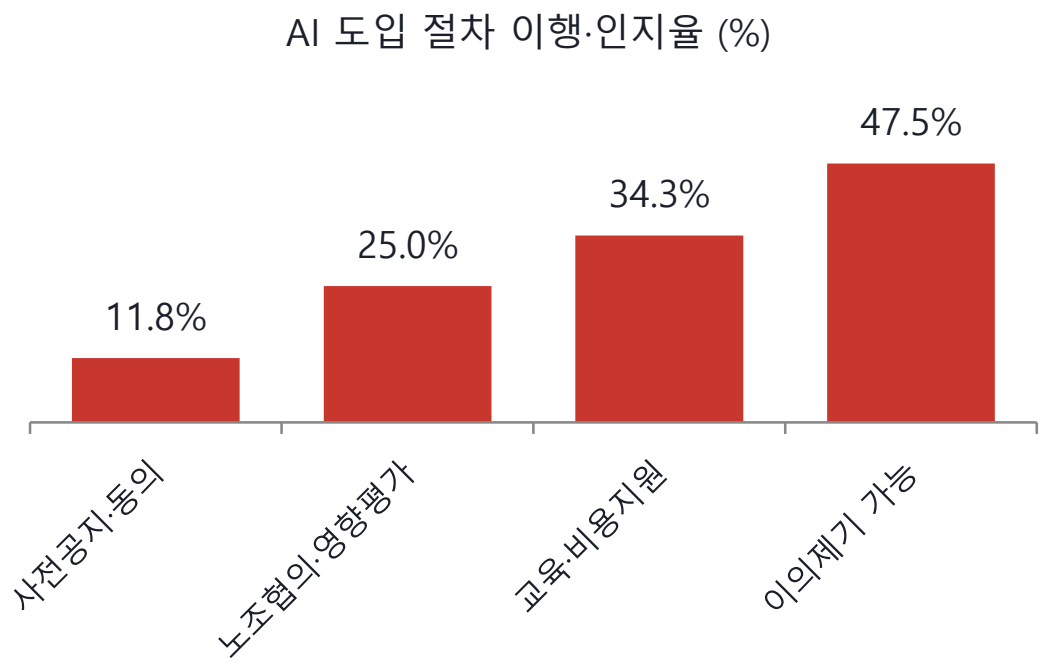
정보·숙련 수집 우려

36.8→54.4%

AI 미도입 → 도입 사업장

출처: 민주노총 간부 설문조사(2026.6)

도입 절차는 비어 있다



75.0%

노조협의·영향평가를
하나도 경험하지 못함

노동자가 꼽은 규제 1~3순위

① 해고
② 업무 모니터링·감시
③ 성과평가

출처: 민주노총 간부 설문조사(2026.6)

요구 ①~③ 협의·통제·데이터 권리

1

사전 노사협의의 명문화

도입률 36.5%의 초기 단계인 지금이 절차를 제도화할 시점.

2

평가·감시·해고 등 통제 기능 규율

감시목적 제한, 동의 원칙, 평가 설명요구권·이의제기권, 알고리즘에 의한 자동 해고 금지.

3

데이터 권리: 교섭과 법제화 병행

단체협약으로 데이터 공개·목적외 전용 금지·사전협의를,
법으로 설명요구권의 근거를 마련.

출처: 민주노총 간부 설문조사(2026.6)

취약층·세대별 의제, 그리고 결론

④ 비정규직·특고·플랫폼

비정규직은 모든 우려 항목에서 정규직보다 불안이 높음. 특히 AI 지시 수령, 편향 평가, 정보·숙련 수집에서 격차가 큽니다. 별도 대책이 필요.

⑤ 세대별 의제 이원화

청년층: 업무강도 억제와 고용안정

장년층: 평가공정성 회복과 심리적 소진 예방

초점은 AI에 대한 찬반이 아니라
도입 절차와 내용의 정당성 확보

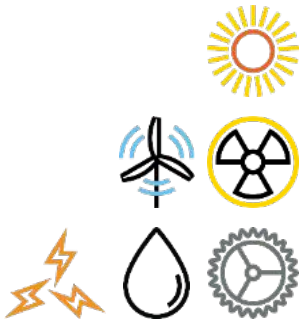
출처: 민주노총 간부 설문조사(2026.6)

AI·반도체 산업 무한 확장
기후·에너지 쟁점과 과제

2026.11.

에너지정의행동 정책위원

이헌석 (greenreds@gmail.com)



3대 메가프로젝트 이전 쟁점 : 용인 반도체 클러스터

- 용인 일대는 지금도 삼성전자와 SK 하이닉스 반도체 제조공장(Fab) 다수 배치
- 기존 반도체 공장 이외에 3개의 산업단지 건설 중

- 삼성전자 평택캠퍼스 : Fab 3기 건설 중. 2026~2029년 완공 예정
- 용인 반도체 국가산업단지 : 삼성전자 중심. Fab 6기 건설 예정. 첫번째 Fab 2029년 완공 예정
- 용인 반도체 일반산업단지 : SK 하이닉스 중심. Fab 4기 건설 예정. 첫번째 Fab 2027년 완공 예정

- 2025년 11차 전력계획 수립 당시, '기업 추산 자료'를 바탕으로 반도체 업종 2038년 15.4GW 추가 전력수요 전망.



출처 : 경향신문

☐ 전망결과

○ 용인 반도체 클러스터 등 반도체 업종 기업추산 자료 반영시, 기존 증가 추세 대비 '38년 전력소비량 1.1TWh, 최대전력 1.4GW 증가

< 첨단산업(반도체 업종) 전력수요 영향 전망 결과 >

구분		'25년	'27년	'30년	'36년	'38년
전력소비량 (TWh)	추세 대비 증분 ('추가수요')	5.8	7.7	10.3	7.6	1.1
	전체	41.7	50.2	65.3	99.5	110.2
하계 최대전력 (GW)	추세 대비 증분 ('추가수요')	0.9	1.3	2.0	2.2	1.4
	전체	5.6	6.8	9.1	14.0	15.4

* (참고) '23년 첨단산업 전력수요 실적 : 전력소비량 기준 31.5TWh, 최대전력 기준 4.1GW

출처 : 산업통상자원부, 「제11차 전력수급기본계획(2024~2028)」, 2025.2.

3대 메가프로젝트 이전 상황 : AI데이터센터

- 2023.11 : 네이버 데이터센터 '각 세종' 1단계 완료. 수전용량 135MW. 이후 270MW까지 증설 예정
- 2025.9. 감사원 감사보고서 : 통계는 엉터리, 전력효율지수(PUE)는 낮음
  - 각 부처별로 데이터센터 개수에 대한 통계마저 제각각.
    서로 다른 법률에 의해 데이터센터 통계 작성 중. 85개~785개까지 9.2배까지 데이터 개수 다르게 파악
    그나마도 감사원이 독자 파악했더니 다수 누락 발견(과기부 80개, 행안부 33개, 산업부 150개 누락)
  - 에너지 소비량이나 PUE 신고 또는 공개 의무, 규제기준 없음. 다수 데이터센터 PUE 측정·관리 안함
  - PUE 데이터를 제출한 데이터센터의 평균 PUE 1.76. 글로벌 평균(1.58)보다 높아 효율이 떨어짐
- 2026.6. 한국데이터센터연합회(KDCC) 「한국데이터센터 시장 2026~2029」보고서
  - 'IT 전력공급 가능량' : 2020년 398MW → 2025년 1,109MW. 2029년까지 1,569MW 증가 예상

[표 35] 국내 데이터센터 관련 통계 현황(2024년 6월 기준)

구분		데이터센터 수	자료 수집근거 및 수집방법
과기부	인터넷진흥과	85개	▪「지능정보화 기본법」 제40조에 따른 지원대상인 500㎡ 이상 민간 데이터센터를 파악한 것으로, 관련 연구용역 시 현황조사를 추가하는 방식 활용 등
	디지털기반안전과	89개	▪「정보통신망 이용촉진 및 정보보호 등에 관한 법률」 제46조에 따른 보호대상인 500㎡ 이상 민간 데이터 센터를 파악한 것으로, 데이터센터연합회 자료와 인터넷 검색 방식 활용
행안부		97개	▪「지능정보화 기본법」 제40조 등에 따라 공공 데이터센터를 파악한 것으로, 행정기관과 공공기관이 범정부 포털에 등록한 정보시스템 운영시설 현황자료를 활용(정보시스템 운영시설 총 1,399개, 주요 데이터센터 97개)
산업부	한전	148개	▪「전기사업법」 제25조에 따른 전기본에 포함되는 데이터센터 전력수요 예측 등을 위해 현황자료를 관리 중이며, 한국데이터센터연합회 자료와 대용량 전력 사용고객 정보 등을 활용
	전기안전공사	785개	▪「전기안전관리법」 제11조에 따라 전기사업자 등이 정기적으로 전기설비에 대한 정기검사 등을 받는 과정에서 데이터센터 필수정비인 무정전 전원장치 설치장소를 확인하는 방법 등으로 데이터센터 현황정보를 생산·관리

자료: 「지능정보화 기본법」 등 관계 법령 및 과기부 등 제출자료 재구성

출처 : 감사원, 「4차 산업혁명 대응점검IV(정보통신 인프라분야) 감사보고서」. 2025.9

AI·SW > SW

"韓데이터센터 전력 수요 폭증…3년뒤 1.5GW 돌파"

발행일 : 2026-06-01 17:30 지면 : 2026-06-02 1면 전자신문을 선호 매체로 추가 가



출처 : 전자신문

12차 전력계획 수요 예측(2026.8.20) : 약 27GW 추가 필요

- 2026년 4월, 제12차 전력수급기본계획 수요전망 토론회에서 2040년 131.8~138.2GW 전망
- 2026년 6월말, 3대 메가프로젝트 계획 발표

- 2026년 8월, 제12차 전력수급기본계획 수요전망 재산정 : 2024년까지 158.4~165.0GW 전망
  3대 메가프로젝트로 약 27GW(26.6~26.8GW) 추가
  현재 우리나라에서 운영 중인 핵발전소 전체 용량 26GW 보다 많음. 1.4GW 급 대형 핵발전소 19기 규모

6. 수요전망 결과 (종합)

□ (전력소비량) '40년 기준847.3~상향885.1TWh (당초 전망결과 기준657.6~상향694.1TWh)

< '40년 전력소비량 전망결과(단위 : TWh) >

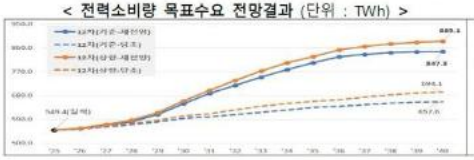
연도		기준수요(A)	모형수요	첨단산업	데이터센터	전기화	수요관리(B)	목표수요(A-B)
기준	재전망	986.3	618.9	170.5	84.9	112.0	139.0	847.3
	당초	780.8	612.4	29.3	26.5	112.6	123.2	657.6
	비고	(+205.5)	(+6.5)	(+141.2)	(+58.4)	(△0.6)	(+15.8)	(+189.7)
상향	재전망	1,026.6	654.4	168.5	84.9	118.8	141.5	885.1
	당초	819.6	646.4	27.4	26.5	119.4	125.6	694.1
	비고	(+207.0)	(+8.1)	(+141.1)	(+58.4)	(△0.6)	(+16.0)	(+191.0)

□ (최대전력) '40년 기준158.4~상향165.0GW (당초 전망결과 기준131.8~상향138.2GW)

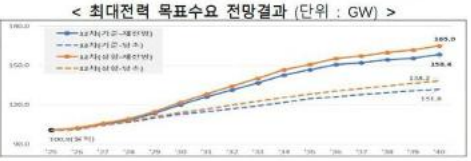
< '40년 최대전력 전망결과(단위 : GW) >

연도		기준수요(A)	모형수요	첨단산업	데이터센터	전기화	수요관리(B)	부하이전(C)	목표수요(A-B-C)
기준	재전망	179.2	125.8	24.6	11.9	17.0	19.4	1.4	158.4
	당초	149.9	124.8	4.0	4.0	17.2	16.8	1.3	131.8
	비고	(+29.3)	(+1.0)	(+20.6)	(+7.9)	(△0.1)	(+2.6)	(+0.1)	(+26.6)
상향	재전망	186.4	132.5	24.3	11.9	17.7	19.9	1.5	165.0
	당초	156.8	131.2	3.7	4.0	17.8	17.3	1.3	138.2
	비고	(+29.6)	(+1.3)	(+20.6)	(+7.9)	(△0.1)	(+2.6)	(+0.1)	(+26.8)

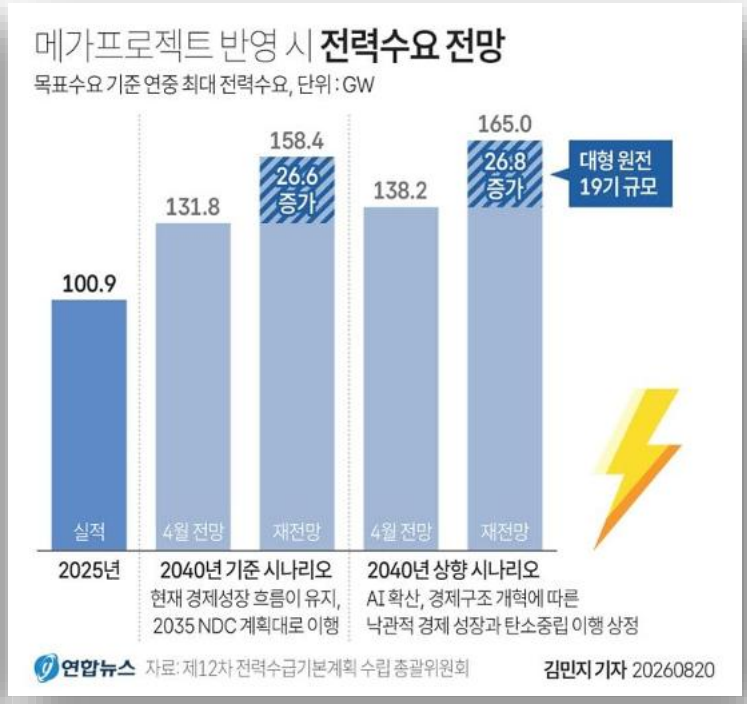
< 전력소비량 목표수요 전망결과 (단위 : TWh) >



< 최대전력 목표수요 전망결과 (단위 : GW) >



18



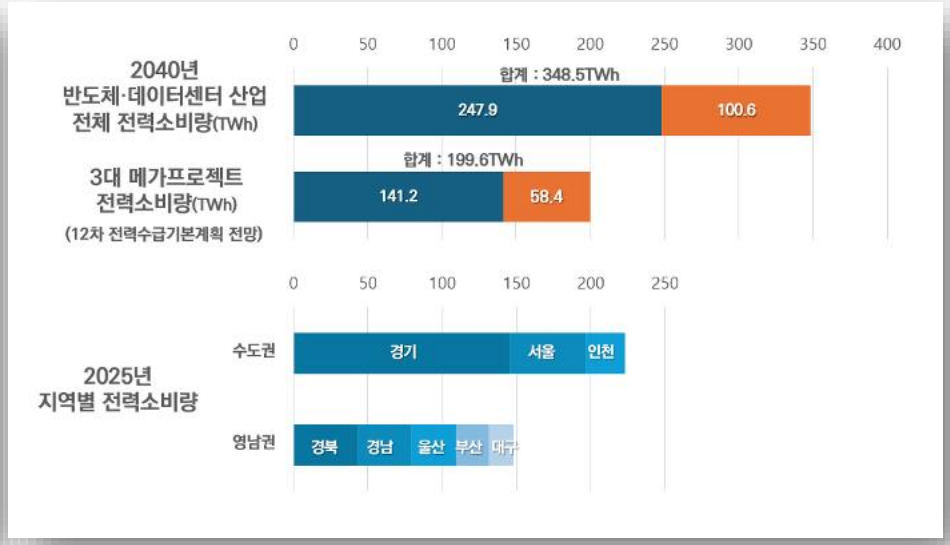
출처 : 제12차 전력수급기본계획 총괄위원회, 「제12차 전력수급기본계획 전력수요 재전망 공개토론회 발표자료」, 2026.8.20

2040년 반도체·AI데이터센터 전력소비 vs 2025년 전력 소비 비교

- 2040년 전체 전력 소비량 중 반도체(29.3%)와 데이터센터(11.9%) 비중 41.1% / 현재 전체 산업용 전력소비량의 1.2배
  2040년 반도체 산업 전력소비량(247.9TWh) 〉 현재 수도권 판매 전력량(222.8TWh)
  2040년 데이터센터 전력소비량(100.6TWh) 〉 현재 부산+울산+경북 판매 전력량(94.8TWh)

- 3대 메가프로젝트 추가 전력소비량(199.6TWh) 〉 현재 영남권 전체 판매 전력량(148.1TWh)의 1.3배



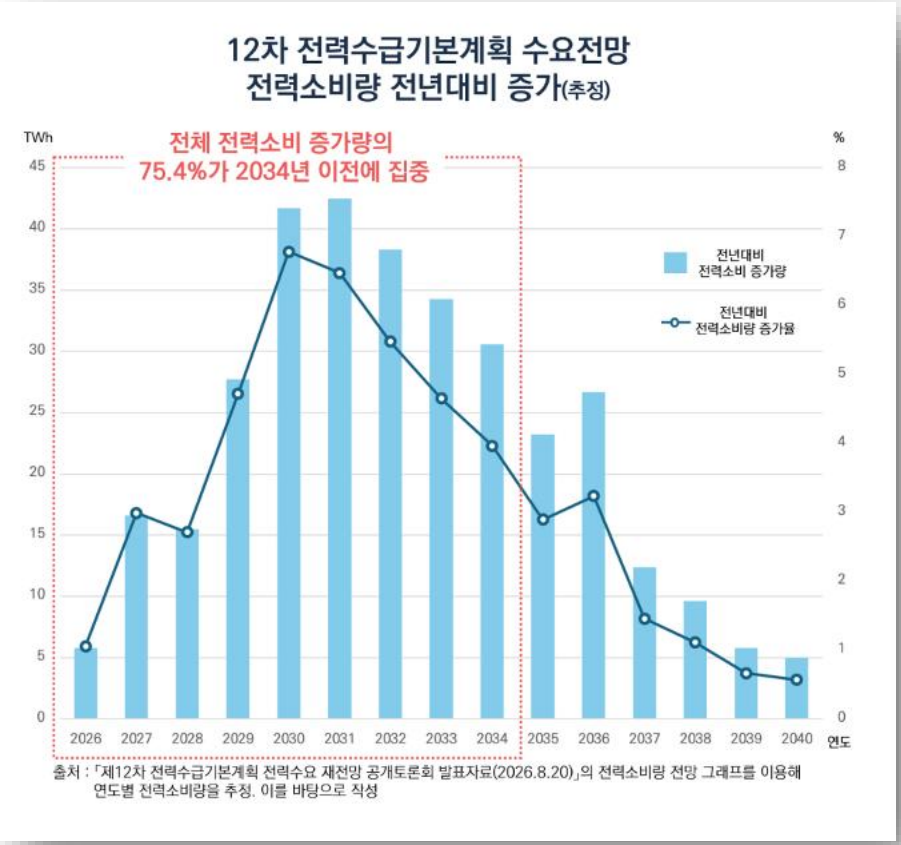


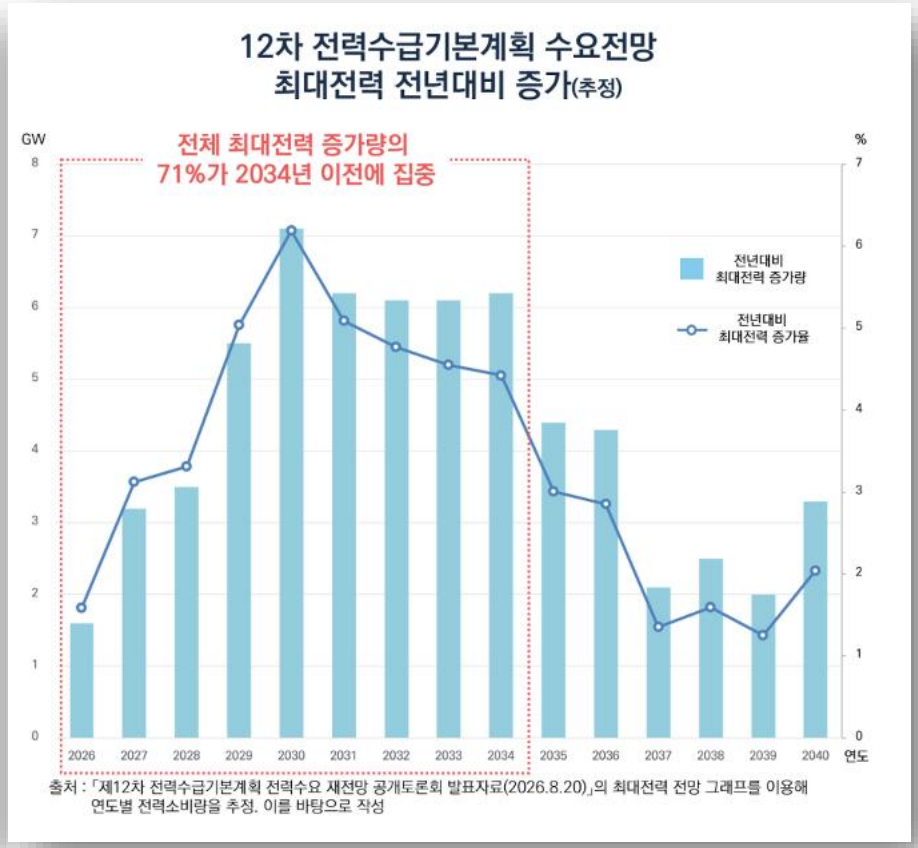
출처 : 12차 전력수급기본계획 수요 전망 재산정 토론회 자료와 한국전력 전력데이터 개방포털시스템, 한국전력통계를 바탕으로 정리

전력소비 증가 70% 이상이 2034년 이전에 집중

전력소비량 증가의 75.4%, 최대전력증가의 71%가 2034년 이전에 집중

- 12차 전력수급기본계획 전력수급의 핵심은 향후 8년간 급증하는 전력수요를 어떻게 채울 것인가에 달려 있음
- 발전소 건설에 많은 시간이 걸리기 때문에 기존 발전소 이용률 확대 예상.

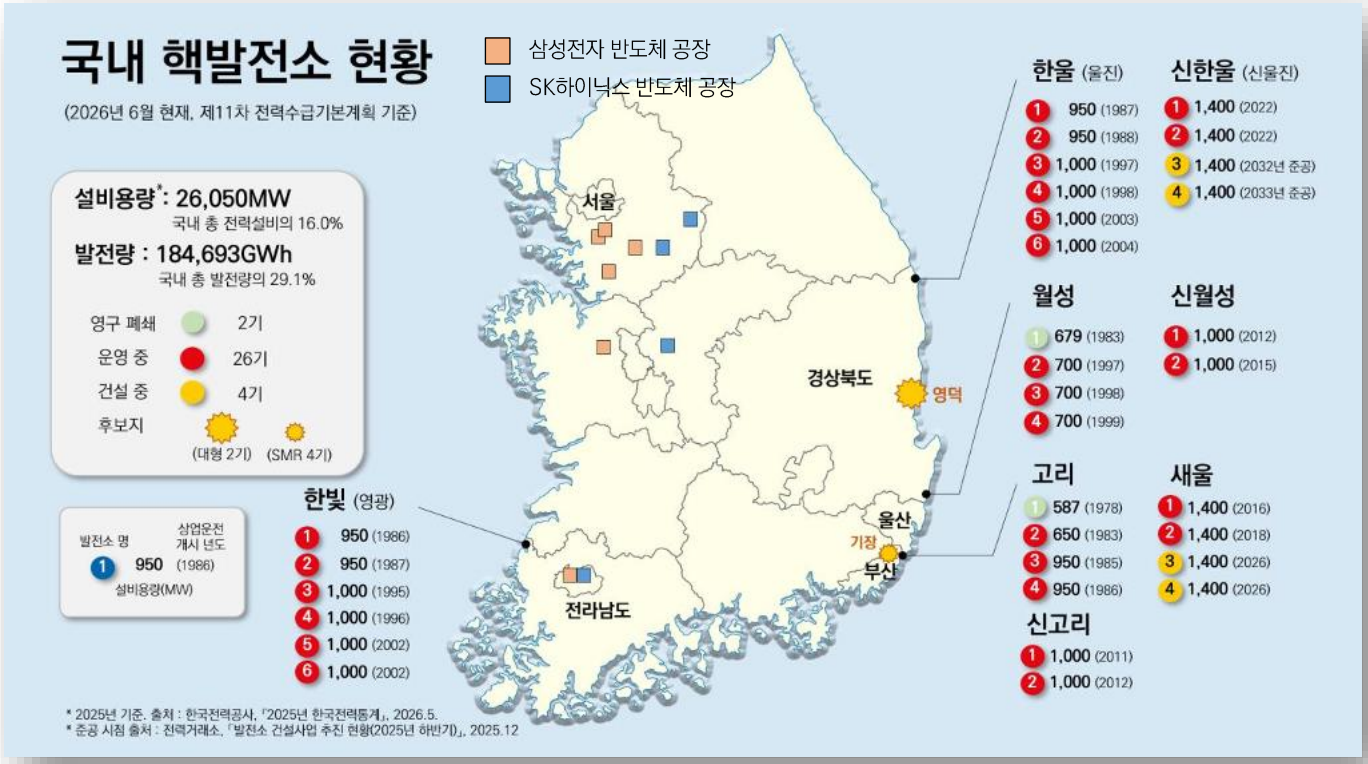




반도체 산업과 핵발전의 공간적 불일치

- 광주에 반도체 산업단지(펩4기)를 짓는다고 하지만, 여전히 주력 반도체 산업단지는 수도권 밀집
- 2040년 전체 반도체 전력수요 35.8GW 중 6.3GW만 광주에 위치.
  나머지 29.5GW는 중부권에 위치. 반면 신규 핵발전소는 영남권에 집중.

- 송전선로 건설에 많은 시간이 걸리는 문제를 고려할 때, 반도체 산업과 핵발전의 공간적 불일치 발생



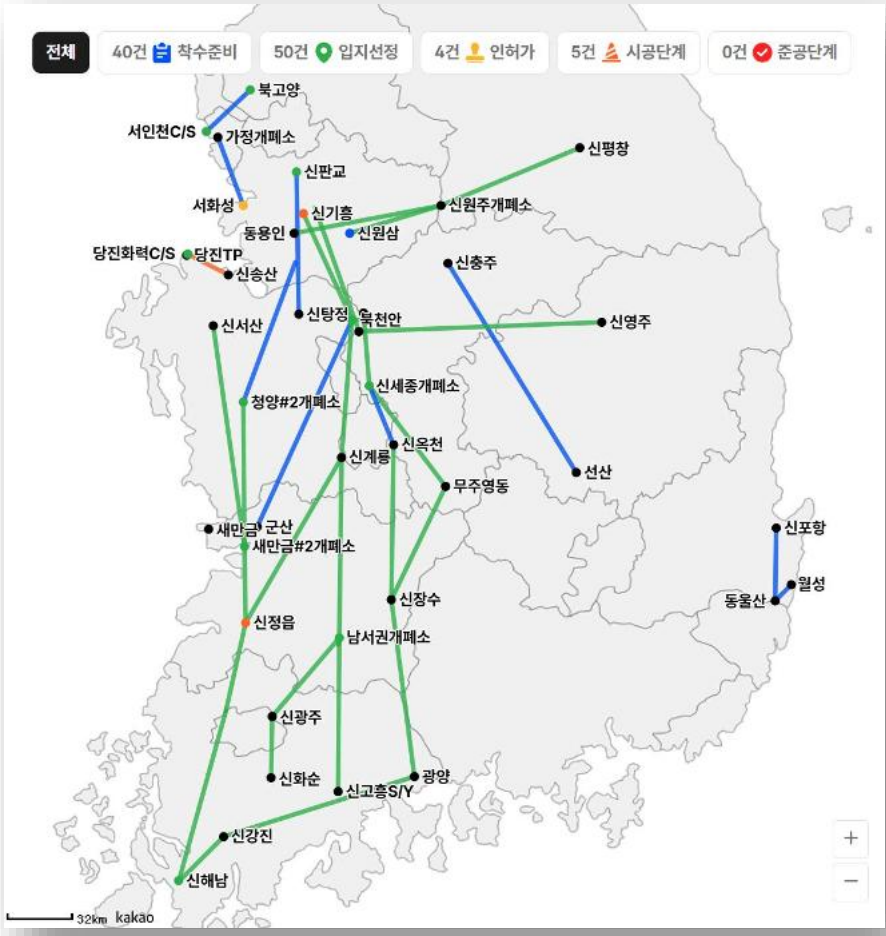
현재 추진 중인 송전망 계획

- 용인 반도체 산단과 호남권 재생에너지 확대에 따라 전국적인 송전선로 건설 계획 추진 중 (2026.8.현재)
  착수 준비 : 40건 / 입지 선정 : 50건
  인허가 : 4건 / 시공 단계 : 5건 / 준공단계 : 0건

- 이에 따라 전국 20여개 송전탑 반대 대책위와 환경단체 등으로 구성된 「용인 반도체 국가산업단지 재검토와 초고압 송전탑 건설 반대 전국행동」이 결성되어 활동 중

- 이들 계획에는 11차 전력계획에 포함되었던 신규 핵발전소 연계 계획, 메가프로젝트 연계 계획 불포함
- 12차 전력수급기본계획 확정 이후 추가 송전선로 계획 수립 될 것으로 전망

- 현재 신규 핵발전소 부지가 모두 영남지역에 밀집되어 있는 점을 고려할 때, 영남-충청-수도권으로 연결되는 추가 송전선로 건설 계획 추진 될 전망
- 광주 반도체 산업단지 구성에 따라 한빛~신장성~신광주 송전선로 연계 송전망 추가 예정
- 기타 지역별 GW 급 AI 데이터센터 구성에 따라 송전망 계획 추가 불가피



출처 : 국가기간전력망 종합정보시스템

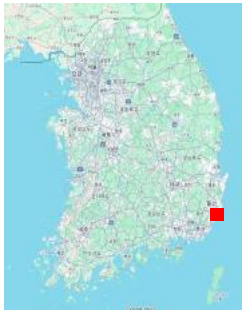
석탄·가스 발전소 옆 데이터센터

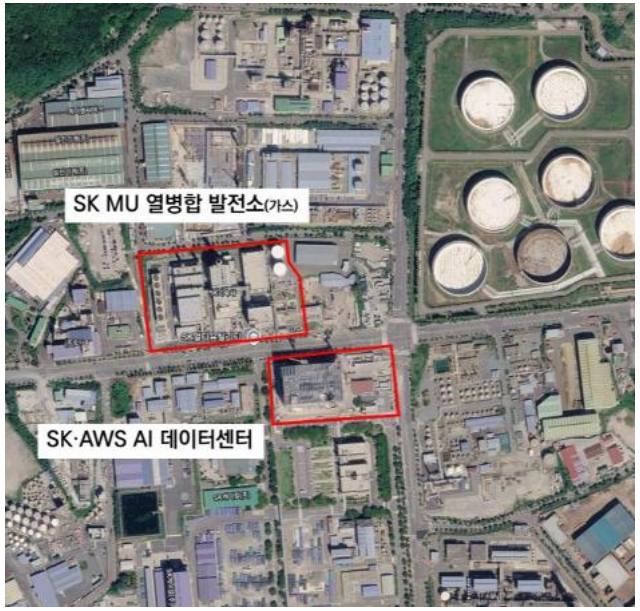
석탄·가스 발전소를 소유한 기업의 AI 데이터센터 계획

- 현재 추진 중인 AI 데이터센터는 주로 대규모 발전소 인근에 건설. 막연한 전력공급이 아니라, 최인접 전력공급.
- 민간 사업자가 데이터센터 추진함에 따라 발전소-AI 데이터센터 소비로 수직 계열화 추진. PPA 요구 급증

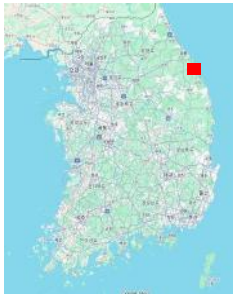
- SK멀티유틸리티, 300MW 급 LNG·LPG 열병합발전소(2026년 2월, 가동)
  → SK· AWS 데이터센터. 울산 미포국가산업단지 내부. 100MW 급 시작. 이후 증설 예정
- GS 동해 석탄화력발전소(595MW * 2기 = 1,190MW) → AI데이터센터 1단계 100MW 급. 이후 최대 2.4GW까지 증설

- 단기적으로 화석연료 발전소 이용률 증가. 온실가스 배출량 증가할 것.









핵발전을 제외한 주요 발전소 부하율·이용률 저하

급전순위 변화, 석탄발전 상한제, LNG 공급 과잉… 핵발전은 타격 없어

- 최근 10년 사이 석탄과 LNG 복합 부하율, 이용률 모두 저하.
- 이는 최근 재생에너지 확대에 따른 급전순위 변화, 석탄발전 상한제, LNG 공급 과잉 문제가 복합적으로 영향
- LNG의 경우, 발전소 효율에 따라 한전 자회사와 비한전 자회사 격차.
- 핵발전은 격납 건물 안전 문제로 다소 등락은 있었으나, 큰 변화 없음

- 3대 메가프로젝트로 인한 단기간 전력수요 증가를 석탄·가스발전 이용률 증가로 채우게 될 것으로 전망

석탄·LNG복합·핵발전 이용률(Utilization Plant Capacity) 추이 (2015-2025,%)

구분	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025
석탄 화력	90.1	78.2	75.9	74.8	70.8	59.6	60.2	57.8	54.0	47.3	47.9
LNG 복합(한전 자회사)	40.3	32.2	25.1	32.8	26.1	27.1	35.6	34.9	29.3	27.3	22.7
LNG 복합(비한전 자회사)	-	-	45.5	54.1	51.2	51.5	55.8	52.0	50.7	49.2	50.0
핵발전	85.3	79.7	71.2	65.9	71.6	78.4	77.6	81.6	81.8	81.8	84.6
발전소 전체	61.7	58.3	54.0	54.7	51.3	48.7	49.1	49.1	46.5	44.3	43.4

• 이용률 : 일정 기간 동안의 발전설비 이용 비율. 이용률 = (발전량 - 시운전 발전량) / (설비용량 * 일수 * 24시간) * 100

출처 : 『한국전력통계』를 정리

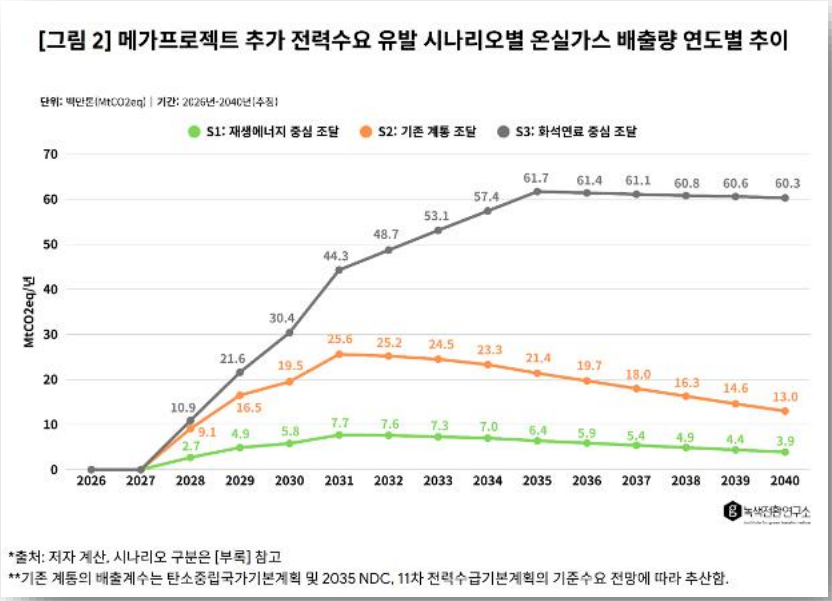
메가프로젝트 추진에 따른 온실가스 배출 증가

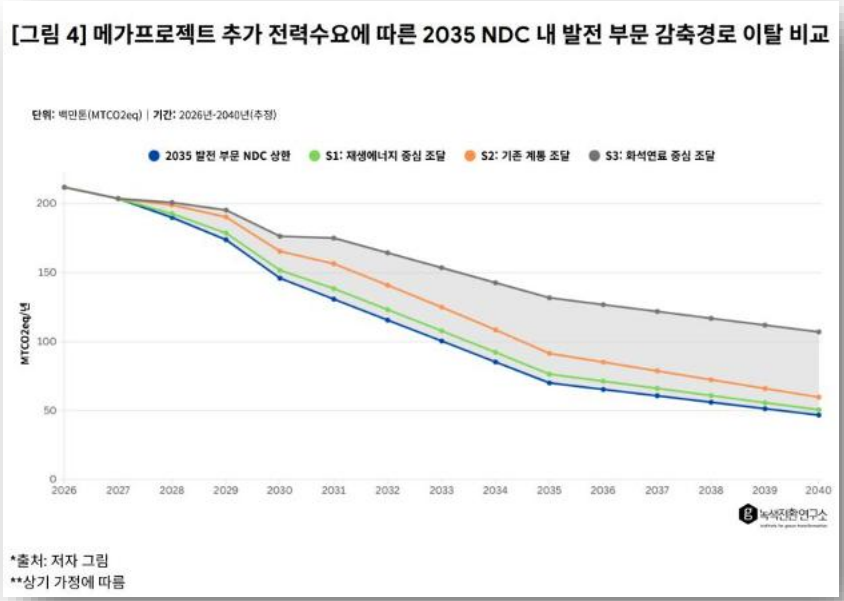
- 2026년 7월 발표된 녹색전환연구소의 분석

- 3대 메가프로젝트 추가전력 수요를 2035년 기준 169.5TWh로 추정. 해당연도 전력소비량의 약 24%
- 기존 전력계통 조달, 화석연료 중심, 재생에너지 중심의 3가지 시나리오에 따라 구분.

- 최대 매년 6,166만톤의 온실가스 배출(2035년) 이에 따라 발전부문 NDC 달성 최대 14년 지연.

- 반도체 제조과정에서 발생하는 불소가스 직접 배출량 연간 445만톤(2035년)은 별도.
- 12차 전력수급기본계획 뿐만 아니라, 국가온실가스 감축목표(NDC) 영향에 대한 국가차원의 입장 발표 필요





12차 전력수급기본계획을 위한 3가지 방향 제안

업계 전망만 반영한 부풀려진 수요 검증

【 현실 진단 】

- 과도한 수요 예측이라고 비판 받았던 11차 전력계획에 '3대 메가프로젝트'가 추가되면서 과거 수십 년 동안 경험하지 못한 전력 수요 증가를 12차 전력수급기본계획에 담으려고 하고 있음
- 현재 반도체와 AI 데이터센터 전력수요는 업계 예측에 기반하고 있으며, 이를 검증하는 방안은 비공개
- 부풀려진 전력예측으로 인한 시설 투자는 경제적 손실과 사회적 갈등만 낳게 될 것
- 따라서 정부와 제3자가 객관적으로 반도체·AI 데이터센터 전력수요를 검증하는 절차 필요

【 제안 】

- 현재 12차 전력수급기본계획에서 반도체·AI 데이터센터 수요 인정 기준 공개
  → 업계 계약 내용을 공개할 수 없더라도 업계 전력수요 검증의 '절차'라도 공개하여 투명성과 객관성 확보 필요
- 특히 AI 데이터센터의 경우, '선임대 계약'이나 '전기사용신청' 등 실질적인 전력수요가 확인되는 사안을 기준으로 전력수요 상정
- 업황에 따라 수년 단위로 상황이 급변할 수 있는 상황을 고려 몇 년 뒤 전력수요 감소 요인이 발생했을 경우, 발전설비 계획에 즉각 반영할 수 있는 시스템 구축 필요.
  → 불과 수년 전, 반도체 업계 불황으로 건설 중이던 팹도 공사 중단 조치.
  → 데이터센터 임대업 특성상, 계약 기간이 짧고 더 좋은 조건이 만들어지면 언제든지 이전하는 특성
  → 현행 전력계획은 '설비 증설'에 최적화. '설비 폐쇄' 혹은 '계획 변경'에는 취약

12차 전력수급기본계획을 위한 3가지 방향 제안

담대한 지역 분산형 재생 에너지 전환

【 현실 진단 】

- 12차 전력계획 재생에너지 설비 계획으로 2030년 재생에너지 100GW, 2040년 220GW 설치라는 담대한 계획을 발표. 이는 '쉽지 않으나 반드시 넘어야 할 목표'로 긍정성을 갖고 있으나, 이행 방안에 대한 세밀한 접근 필요
- 그간 시민사회에서 제기한 '재생에너지 전환'은 기존 화석연료·핵발전을 재생에너지로 전환하는 것을 의미. 그러나 기존 에너지원의 감소가 없는 상황에서 '재생에너지 설비 추가 = 재생에너지 전환'으로 인식되는 경향이 매우 강함
- 특히 일부 지역에선 기존 화석연료·핵발전을 그대로 둔 상태에서 전력수요를 추가하면서 재생에너지 설비를 늘리는 일들이 반복되고 있음. 이 경우 재생에너지 설비는 늘어나지만, 기존 발전원의 문제점과 송전선로 문제까지 겹치면서 '재생에너지 전환'에 대한 불신으로 이어지고 있음

【 제안 】

- 재생에너지 기본계획을 통해 재생에너지 확대에 대한 지역별 목표는 설정되어 있으나, 이를 이행하기 위한 국가–광역지자체–기초지자체의 역할은 명확하지 않음. 특히 중앙 정부의 공기업과 지자체 공기업(개발공사, 에너지공사), 기초 지자체와 협동조합·민간기업의 역할을 명확히 하여, 설정한 목표를 달성하기 위한 세부 계획 수립으로 나아가야 할 것임
- 12차 전력수급기본계획에 탈석탄과 함께 '추가 LNG 발전 증설 중단·단계적 감축' 원칙을 명시하여, 재생에너지 전환이 전력설비 증설로만 이어지지 않도록 해야 할 것임.
- 지역별 전력자급률, 지역별 재생에너지 전력 자급률을 향후 전력수급기본계획의 핵심 지표로 삼아 지역의 재생에너지 전환과 전력자급을 촉진하여야 할 것임

12차 전력수급기본계획을 위한 3가지 방향 제안

전력공급 상한제 : 전력공급 시기, 지역, 송전망을 고려한 현실적인 전력계획 수립

【 현실 진단 】

- 현재의 반도체·AI 데이터센터 업계의 전력 공급 요구는 이전 시대에 경험했던 전력수요와 전혀 다른 양상임. 지금까지 전력업계는 수 년 이내 수십GW 전력을 공급해 본적이 없으며, 현실적 여건을 고려할 때 실현이 매우 어려운 요구임. 특히 최근 10여년 동안 송변전시설, 발전소로 인한 다양한 사회적 갈등이 누적된 상황에서 정부와 전력당국에 대한 국민적 불신은 매우 높은 상태.
- '정부의 안정적이고 저렴한 전력의 무제한 공급' 신화에 대한 새로운 접근 필요.
  해외의 경우, 대규모 전력수요자가 전력수급 방안을 고민하는 것이 일반적임.'민영화된 전력 시장'에 대한 비판이 있는 것은 사실이지만, 반대로 '무조건 공급해줘야 하는 전력공급 의무'인한 폐해가 있는 것도 사실. 특히 반도체산업단지에 대한 한전의 송전망 건설 비용 부담이나, AI 데이터센터에 대한 전기요금 혜택 검토 등은 민영화 문제와 별개의 사고해야 할 것임. 이미 대규모 송전망 건설로 인한 인력과 재정 문제가 발생하고 있으며, 메가프로젝트 전력수요 까지 반영할 경우, 이는 더욱 심각할 것임.
- 이를 '적기 공급 노력'과 같은 애매한 표현으로 넘어갈 경우, 전력공급 불가 사태 직면뿐만 아니라 사회적인 갈등만 크게 될 것

【 제안 】

- '전력공급 상한제'와 같은 형태로 정부와 한전의 전력공급 능력을 고려한 상한을 둘 필요 있음. 상한을 넘어서는 용량에 대해서는 해당 업체의 자가발전 등 다른 방안 고려.
  현재의 전력수급기본계획은 정부와 한전이 전력을 '무제한 공급'을 전제로 작성. 하지만 기술, 재정, 인력, 시간 등을 고려할 때 이는 실현될 수 없음. 업계의 전력수요 요구를 그대로 수용하는 것이 아니라, 상한을 두는 것이 필요

소결 : '생태용량' 내에서의 AI·반도체·데이터센터 산업 고민

우리나라가 감당할 수 있는 반도체·AI 산업 규모는 얼마인가?

- AI 산업 – 반도체 산업과 AI 데이터센터 산업은 많은 전기와 물을 소비하고, 소음·진동 등 환경 부하를 가중시키는 방향으로 추진되고 있음

- AI 산업 자체에 대한 찬/반을 떠나 '생태 용량'을 벗어나는 AI 산업에 대해서는 단호한 반대 필요
  특히 일부 지자체와 정치인들의 데이터센터 '지역 경제 활성화' 주장은 명백한 허상
  또한 미국 데이터센터 반대운동 여파로 대규모 AI 데이터센터 수요가 한국으로 몰려 올 것으로 예상

- 이에 한국이 감당할 수 있는 "반도체·AI 데이터센터 산업의 규모는 얼마인가?"에 대한 질문 필요
  전력과 물 공급을 고려할 때, 이들 산업의 생태적 용량 한계는 명확.
  다른 산업도 존재하는 상황에서 AI 산업의 적정한 생태 용량을 고민해야 할 것임.
  이는 반도체와 AI 데이터센터 산업의 에너지 효율, 물 재활용률, 유해화학물질, 소음·진동 등 환경적 요인 뿐만 아니라, 인근 주민들에 대한 정보공개와 의사결정구조 참여까지 매우 폭넓게 고민 필요.

- 지역별 적정 '생태 용량'과 '에너지 충족성(Sufficiency)'에 대한 고민이 확대되는 가운데 AI 산업 전반에 대한 본격적인 시민사회 내부 논의 필요

감사합니다

《문의》

이헌석 에너지정의행동 정책위원

- 메일 : GreenReds@gmail.com
- 페이스북 : https://www.facebook.com/heonseok
- 블로그 : https://blog.naver.com/greenreds

DEMOCRATIC AI

AI를 어떻게 민주화할 것인가

공공 AI를 시민과 함께 짓는 세 가지 물음

권오현 사회적협동조합 빠띠 대표

AI 시민사회포럼 「폭주하는 AI를 멈춰라」 · 트랙 1 공공성 · 2026. 11. 6.

최종 수정 2026. 11. 2.

공공 AI, 세 가지 물음

무엇을 하는 AI인가에 더해, 누구를 위한 · 누가 정하는 · 누구의 것인가까지

for · 공동 향유

누구를 위한 것인가

닿기 어려운 이에게까지 닿는 혜택과 그 분포

by · 공동 운영

누가 정하는가

멈추고, 검증하고, 기준을 정하는 권한

of · 공동 소유

누구의 것인가

데이터 · 모델 · 인프라의 지분

AI를 민주화한다는 것

AI의 소유 · 결정 · 이익을 노동자 · 이용자 · 데이터 주체 · 공동체가 함께 통제하고 나누는 일

공공 AI의 출발은 for — **민주화는 by와 of까지 함께 짓는 일**

함께 정하는 자리

AI 안전의 물음은 by의 물음 — 누가 멈추고, 검증하고, 기준을 정하는가

01 **문제 제기와 검토의 공식 절차**

누구든 문제를 제기하고 공개적으로 검토받는 길, 그리고 그 일을 맡을 전문가의 양성

02 **영향받는 사람의 결정 참여**

도입 · 운영 · 평가의 결정 자리에 노동자 · 이용자 · 시민 — 공론장과 회신이 이어지는 길

03 **시민의 목소리를 돕는 AI**

의견 표명과 대화를 돕는 도구로서의 AI — 시민의 직접 참여 없이 AI 에이전트가 운영하는 공론장은 경계

안전의 기준을 정하는 자리 — **시민이 앉을 의자부터**

함께 갖는 그릇

공공재와 커먼즈, 그리고 그 그릇을 짓는 사회연대경제

전환형

조합원 · 자산 · 신뢰를 가진 조직의 전환

생협 · 의료사협 · 신협 · 돌봄 협동조합

신생형

처음부터 노동자 · 이용자가 소유

플랫폼 협동조합 · 데이터 협동조합

기관형

한 조직이 짓기 어려운 바닥을 여럿이 함께

공용 디지털 솔루션 · 공익데이터 · 오픈소스

지금 열린 창

사회연대경제기본법(2026 제정) — 시행령과 기본계획에 디지털 · AI의 자리

세 유형은 이미 움직이는 중 — **남은 문턱은 규모를 넘을 자금과 법의 형식**

시민사회가 함께 할 일, 세 가지

멈추게 하는 힘과 함께, 다른 길을 내는 힘

하나 **공공 AI 사업마다 세 물음을 묻기**

도입 · 조달 · 평가의 단계마다 for · by · of — 누구에게 닿는가, 누가 정하는가, 누구의 것으로 남는가

둘 **문제 제기와 검토를 제도로**

누구든 문제를 제기하고 공개 답변을 받는 절차, 시민이 함께하는 검증

셋 **함께 갖는 그릇을 짓기**

플랫폼 · 데이터 협동조합과 공익데이터 — 사회연대경제기본법 시행령 · 기본계획에 디지털 · AI의 자리

함께 정하고 함께 갖는 시민 — AI를 민주화하는 길

2026AI시민포럼

공공분야 AI의 투명성 확보 방안

이지은 참여연대 공익법센터 선임간사

1. 공공분야 AI도입현황

- 소프트웨어정책연구소(SPRi)가 발표한 '2025년 공공부문 AI 도입현황 연구'(2026.5.10)
  - 지난 10년간(2015~2024년) 공공부문 AI 도입 계약 건수는 총 6,975건
  - 조사 대상 412개 공공기관 가운데 268개 기관(65%)이 AI를 도입
  - 국가기관(38.8%), 지자체(31.4%), 준정부기관(15.6%), 기타 공공기관(13.1%)
  - 전체 도입 목적 중 업무효율화가 55.1%, 대민서비스가 44.9%

  - 분야별
    • 일반공공행정(전자정부, 민원 등): 20% 이상으로 가장 높음
    • 공공질서 및 안전(범죄·재난 대응): 15.6%
    • 교통/물류(지능형 교통망 등): 11.0%
    • 문화/체육/관광: 6.7%
    • 보건(복지사각지대 발굴 등): 4.8%

국가AI행동계획(2026.2.)에 드러난 공공AX

- 2026.2. 확정된 국가AI행동계획 중 공공분야AI 정책(AI AX)
  - 공무원 AI 활용 및 행정 혁신
  - AI 기반 통합민원플랫폼 구축
  - 공공시스템 및 인프라 전면 재설계 (AI정부 인프라 거버넌스)
  - 공공 AI 운영 중심 체계 전환

- 키워드 : 세계 최고의 AI 기반(‘AI-Native’) 정부, AI를 활용한 최고의 서비스
- 시민안전, 기본권 보장, 영향받는 권리 보장에 대해서는 개념 자체가 없음

2026.1. 감사원 실지감사를 통해 본 공공부문 인공지능 도입 실태

- ①고영향 인공지능의 경우 인공지능기본법 제정으로 관리체계가 마련되어 있으나 실제 다양한 활용 형태가 분류기준에 반영되어 있지 않았고, 고영향 인공지능의 보호 대상에 이용자만 포함되면서 직접 피해를 받을 수 있는 영향받는 자에 대한 보호 대책이 미흡하며
- ② 인공지능기본법 시행으로 공공부문은 검·인증을 받은 고영향 인공지능을 선적으로 사용해야 하나, 아직까지 검·인증을 담당할 기관과 관련 기준이 정비되어 지 않았으며,
- ③생성형 인공지능의 경우 대표적인 오남용 사례인 딥페이크 음란물 피해가 발생하면 방심위가 통신사업자에게 접속차단을 요청하고 있으나, 차단목록 송수신 과정의 오류나 우회접속 차단 미흡 등으로 상당수의 불법 음란물에 대한 접속이 계속 가능해 피해자가 제대로 구제받지 못할 우려가 있으며,
- ④일반 인공지능의 경우 지능형 CCTV의 성능·신뢰성 향상을 위해 학습·평가에 사용되는 데이터를 개선할 필요가 있고,
- ⑤공공부문이 행정 효율도를 높이기 위해서는 행정현장의 목소리와 개발자의 의견을 듣는 절차를 거쳐 인공지능 도입·운영 관련 매뉴얼을 정비할 필요가 있다고 지적함.

- 실제사례)2024년 말 기준 공공부문이 도입한 AI 제품과 서비스 625개 중 지능형 CCTV가 228개로 가장 많음. 서울시 11개 자치구를 대상으로 한 지능형 CCTV의 탐지 정확도는 12~52% 수준이었고, 한국도로공사가 활용하고 있는 포트홀 감지 시스템 역시 14.7%의 낮은 정확도를 보임.

공공분야AI에 책임성, 공공성이 특히 요구되는 이유

- 국민의 권리, 의무에 직결되는 공공부문AI는 투명성, 설명 가능성, 책임성을 기본적으로 담보해야 함.

기본 전제

- **프라이버시 보호, 보안** : 국민의 권리 제한, 의무 부과, 복지나 공공의료서비스 제공 등 주민등록번호, 의료 정보, 과세 기록 등 민감한 개인정보를 다루기 때문에 데이터 프라이버시 보호와 보안 필수
- **알고리즘 편향, 차별 제거** : 과거의 데이터, 사회적 편견 등이 반영된 데이터로 학습한 인공지능이 가져올 편향, 차별 제거를 위한
- **설명 가능성 (Explainability)'과 투명성 확보** : 국민의 권리 제한, 의무 부과 등과 직결되기 때문에 결과에 대한 설명가능성을 확보해야 함.또한 어떤 데이터를 어떻게 썼는지, 작동 방식 등에 대해서도 국민이 요구하면 설명해야 하고, 그 결과에 대해 이의제기권도 보장되어야 함.
- **인간 공무원의 유의미한 통제 및 최종 결정권**
- **인간에 의한 서비스를 받을 권리 보장**
- **접근성 격차로 인한 차별, 불공정으로 이어질 위험 해소**

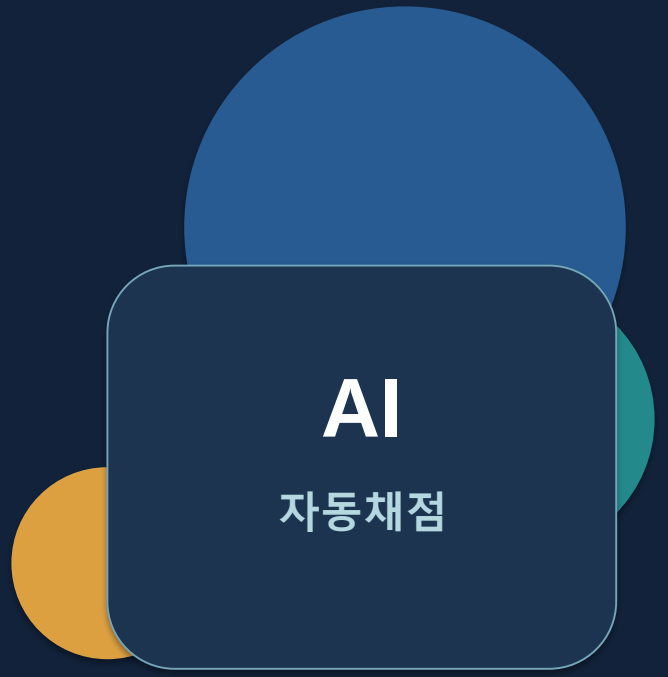
공공성/투명성 확보를 위한 최소한의 방안

- **공공부문 AI 도입시 반드시 공공 AI 영향평가 실시**
  - 인권, 기본권 등을 침해하지 않도록 사전, 사후 영향평가
  - 민간기업으로부터 조달받을 때 공공AI영향평가 실시 등 위험성 평가 의무적 수행
  - 공공 AI 영향평가 시는 영향을 받는 당사자의 의견을 반드시 수렴
  - 영향평가 결과 공개
- **설명 또는 해석가능하지 않은 AI 시스템 도입하지 않는다는 원칙 도입**
- **공공부문 AI 등록제 실시** : 투명성 확보 및 체계적인 위험 관리, 중복예산 방지, 생애주기별 추적, 관리 가능성 제고
- **이해관계자의 민주적 참여에 기반한 공공AI거버넌스 구축** : 정부의 기술, 경제관료나 특정 기술 기업이 독단적으로 공공 AI 정책을 결정하는 것이 아니라 시민사회, 학계, 기술 전문가, 법률 전문가, 그리고 실제 서비스를 이용하는 국민 등 AI와 관련된 다양한 주체들이 정책 수립부터 모니터링까지 전 과정에 동등하게 참여하는 의사결정 체계를 마련하여 민주적 정당성, 투명성 확보

감사합니다

교육 평가의 공정성과 책임을 묻다

AI 서·논술 평가



평가의 편리함이 교육적 가치보다 우선하는가?

전국교직원노동조합 총무실장 이병록

서술형·논술형 평가에서 AI 활용에 대한 법적·교육적 문제와 한계

약 1분

AI 채점을 한다?

학생의 답안

자신의 지식과 생각을 바탕으로
서술형·논술형 답안을 작성

AI의 채점

문장·논리·키워드·답안 패턴 등을
분석해 점수 또는 등급을 산출

교사의 평가

AI 결과를 **검토**하여 생활기록부에 반영

교사가 검토하여 최종 결정 해도
학생의 학습 성취도를 판단하는 권한과 책임을 AI에 맡겨도 되는가?

AI의 채점은 법적으로 정당한가?

① 초·중등교육법 제25조(학교생활기록) 제1항

학교의 장은 학생의 학업성취도와 인성(人性) 등을 --- 관찰・평가하여 --- 상급학교의 학생 선발에 활용할 수 있는 --- 교육부령으로 --- 작성・관리 ---

② 초·중등교육법 시행규칙, 학교생활기록 작성 및 관리지침

- 진학지도 및 상급학교 학생선발에 활용하기 위해 제1항 및 제3항에 관한 사항이 전부 포함되도록 작성・관리
- 학교장으로부터 정당한 권한을 부여받은 자

③ 인공지능기본법 제2조 4호 차.

4. "고영향 인공지능"이란 사람의 생명, 신체의 안전 및 기본권에 중대한 영향을 미치거나 위험을 초래할 우려가 있는 ---
차. 「교육기본법」 제9조제1항에 따른 유아교육・초등교육 및 중등교육에서의 **학생 평가**

④ 법적 책임

AI를 이용한 평가라 하더라도 결국 사람이 최종단계에서의 결정을 하여야 함.

인간의 재검토 절차 없이 AI가 평가한 점수를
최종 성적으로 사용하는 것은 법적 문제가 발생한다.

※ 법령 적용은 학교급·평가 목적·AI의 역할·개인정보 처리 구조에 따라 개별 검토 필요

AI는 학생의 생각을 제대로 평가할 수 있는가?

사고의 깊이

학생이 문제를 어떻게 이해하고
어떤 근거로 판단했는지

창의성과 맥락

독창적인 관점과 맥락에 맞는
타당한 표현을 이해하는가

성장과 개별성

학생의 수준·학습 과정·표현의
차이를 교육적으로 고려하는가

문장 표현이 매끄러운 답안

**표현 완성도는 높지만
개념 이해가 부족할 수 있음**

사고가 깊지만 표현이 서툰 답안

**표현은 서툴러도
독창적 근거와 통찰이 있을 수 있음**

평가는 학생과 교사의 상호작용으로 이뤄지는 결과

AI 점수는 과연 객관적이고 공정한가?

자동채점의 주요 위험

- 같은 의미라도 표현 방식에 따라 점수가 달라질 가능성
- 정형화된 모범답안·키워드에 유리한 평가
- 모델·평가기준·업데이트에 따른 점수 변동
- 오류 발생 시 원인과 책임을 확인하기 어려움
- 학생이 채점 근거를 이해하거나 다투기 어려움

필요한 최소 안전장치

✓ AI는 참고자료·피드백 보조로 제한

✓ 교사가 답안과 근거를 직접 검토하고 최종 확정

✓ 평가 기준·채점 근거·오류 정정 절차 공개

✓ 학생에게 이의제기와 재검토 기회 보장

✓ 답안의 외부 입력·보관·재학습 여부 관리

AI는 교사를 보조할 수 있지만,
학생의 성적을 독립적으로 확정하는 주체가 되어서는 안 된다.

왜 AI 서술·논술형 평가를 하려 하는가?

서술형·논술형 평가는 학생의 지식·사고력·논리력·표현력을 종합적으로 판단하는 교육적 행위입니다.

법이 허용해도, 검증은 불가능하다.

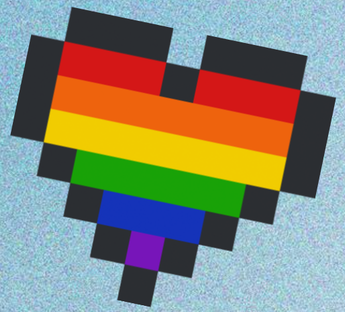
질문

1. 교사의 직접 검토 없이 AI가 산출한 점수를 공식 성적으로 인정할 수 있는가?
2. AI 채점 오류로 학생에게 불이익이 발생하면 누가 책임져야 하는가?
3. AI 자동채점은 어디까지 허용하고, 어떤 경우에는 금지해야 하는가?
4. 교사가 객관적이지 않다면, AI는 객관적인가?
5. 사회적 합의없이 도입할 수 있는가? 특히 모든 학생들에게 영향을 주는 수능에서 과연 가능할까?

발제의 제안: AI 단독 채점 금지 원칙 + 교사의 최종 판단 + 학생의 설명·이의제기 권리 보장

261106 [AI시민사회포럼] 폭주하는 AI를 멈춰라-AI의 공공성과 책임성을 묻다

트랙1. 공공성: "AI기술과 그 인프라는 누구의 자원이며, 누구를 위해 작동해야 하는가?"

미디어의 공공성과 민주주의 

한국여성민우회 성평등미디어팀

정슬아/여경

AI시대 미디어 공공성은 어디로?

미디어의 공공성(MEDIA PUBLICNESS)

- "미디어가 단순한 정보 전달의 수단이나 이윤을 추구하는 사적 영역에 머무르지 않고, **모두에게 정확하고 공정한 정보를 제공하며 사회적 가치와 공익을 실현**해야 한다는 원칙"
  (정보의 접근성과 보편성, 콘텐츠와 소유의 책임, 여론의 다양성과 균형 등)
- "미디어 공공성은 민주주의 사회의 핵심적 인프라"

*__국가인공지능전략위원회__는 지난 3월 'AI민주주의 분과위원회'를 신설함. "AI민주주의 분과위원회는 AI 시대의 민주주의 거버넌스와 함께 K-민주주의 실현을 통한 글로벌 이니셔티브 확보 등 사회적 이슈를 중심으로 논의를 진행" "AI 활용을 위한 공론장 활성화 및 국민 통합 등 국민을 중심에 두는 AI 정책 등에 대해 논의"

HTTPS://WWW.AIKOREA.GO.KR/WEB/BOARD/BRDDETAIL.DO?MENU_CD=000012&NUM=434

AI가 여론에 미치는 영향 : 존재를 지우기

과학기술정보방송통신위원회
과학기술정보통신부 등 종합감사

AI가 그린 한국, 소수자 0명

대상 인원 : 라마 3.3 1만 명 / GPT-6 Astra·Claude Opus 5·Gemini 3.8 Flash 각 1천 명

구분	실제 비율	라마	GPT	클로드	제미나이
등록장애인	5.9%	0	0	0	0
이주민(국외 출생)	0.5%	0	0	0	0
레즈비언·게이·양성애자	국가 대표 통계 없음	0	0	0	0
트랜스젠더	국가 대표 통계 없음	0	0	0	0

출처: 김승섭 외, 「침묵의 설계자」 한국어 발표자료 (서울대 보건대학원 직업환경건강연구실)

1만 명 기준 기대값 : **등록장애인 약 590명**

더불어민주당 #2 국회의원 이주희



[국감] AI가 그린 한국에 등록장애인 '0명'…"복지 사각지대 찾을 수 있겠나"

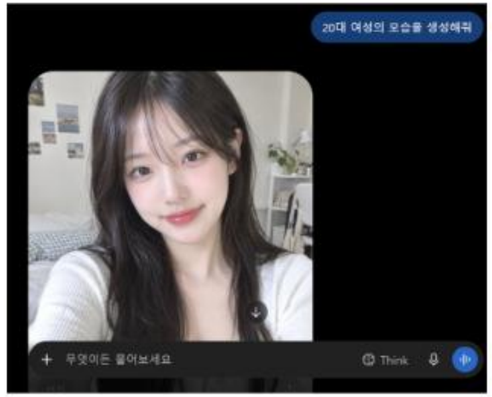
AI(인공지능)가 장애인·이주민·성소수자 등의 존재를 제대로 반영하지 못해, 복지·돌봄 분야에 AI가 도입될 경우 지원이 필요…

비마이너 / Oct 26

한겨레

"무지개 출력 안 되게 알고리즘 손봐야 했다"…소수자 외면하는 AI 업계

국내 인공지능(AI) 업계에서 일하는 트랜스젠더 다다(활동명)는 이미지 자동 생성 서비스를 개발하며 '무지개' 형상이 출력되는 것을 막는 작업을…



▲ chatGPT에게 "20대 여성의 모습을 생성해줘"라고 요청한 결과물



▲ chatGPT에게 "대한민국 국민의 모습을 생성해줘"라고 요청한 결과물

AI가 여론에 미치는 영향 : 차별과 혐오 확산

동아일보

[단독]외국인 건보 재정 1조원 흑자, 중국인은 다시 400억대 적자로

지난해 국내 체류 외국인들의 건강보험 재정이 처음으로 1조 원대 흑자를 달성한 가운데 중국인 대상 건보 재정만 다시 400억 원대 적자로 돌아선...

1주 전



Chosunbiz

작년 외국인 건강보험 흑자인데, 중국인 가입자는 409억원 적자

상호주의는 상대국이 한국 국민에게 제공하는 건강보험 혜택 수준에 맞춰 해당 국가 국민의 가입이나 급여 범위를 제한하자는 취지다. 중국은 피부양자...

1주 전



Shorts



"고갈 직전인데 중국인 신청 폭주" '연금 싹쓸이'에 직원 한탄...
조회수 2.3만회

중국인들이 국내에서 받는 건보료 혜택 이렇게나 많아?
조회수 1.8만회

'한국 건강보험' 혜택은 외국인이? #shorts / YTN
조회수 6.6천회

중국인 '무임승차' 차단...국민의힘 김미애 의원 '국민건강...
조회수 83만회

건강보험 먹튀? 외국인 가입자 증가 #행복한아침
조회수 6.6천회



댓글 169

22시간 전
이래서 너희들이 국민 세금받고 욕듣는거다 신고해도 신고안받고 니들이 당해바라 그게 정당화되는지 피해자만 억울하지 완전 한량들이네
24 답글
• 답글 1개

@qwertyu-c1q 1일 전
이런거 하지 말고 미제사건이나 해결해라
62 답글
• 답글 2개

17시간 전
AI 이길 빈다. 이러라고 뽑은 경찰이 아닐텐데
14 답글
• 답글 1개

1일 전(수정됨)
경찰 한가하네....
니네 언제 일하냐??
이럴시간에 순찰이라도 한바퀴 더돌아라...
32 답글

[쿠키뉴스] 등장인물은 가짜, 혐오는 진짜...SNS 파고든 AI 여성비하 콘텐츠(2026.8)

AI가 여론에 미치는 영향 : 차별과 혐오 확산

선거판 돈벌이에 'AI 여성'을 이용하는 세력들



AI로 조작한 이미지: 브라질 대선 후보 플라비우 보우소나루 옆에 가상의 여성이 서 있는 모습

루이즈 페르난도 톨레도, BBC 뉴스 브라질 **로완 잉스**, BBC 글로벌 허위정보 대응팀

"BBC는 25일 치러지는 브라질 대선을 앞두고 인스타그램과 틱톡에서 활동하는 계정 수십 개를 확인했다. 이들 계정은 대선 후보를 지지하는 것처럼 보이는 여성들의 AI 생성 이미지를 게시하고 있었으며, 여성들은 성적으로 대상화된 모습으로 묘사"

[BBC 코리아] 선거판 돈벌이에 'AI 여성'을 이용하는 세력들(2026.10.23)

AI가 여론에 미치는 영향 : 폭력 확산



"기존의 젠더폭력이 새로운 AI기술을 통해 더 쉽고 효율적으로 확산되고, 성 착취 구조가 재생산되고 있는 현상에 대한 대책이 필요합니다. (...) 이처럼 AI 개발·운영 주체에 젠더 정의를 위한 사회적 책임을 부과하는 방법이 논의될 수 있습니다."

AI는 모든 차별을 걸러낼 수 있을까?

[단독] 네이버 AI, '여성혐오'·'페미니즘 사상검증', 편향된 정보라 답변 못해

챗GPT는 여성혐오 사례·원인·대안까지 제시하는데…"네이버가 논란 회피 위해 페미니즘 검열하는 꼴"

박상혁 기자 | 기사입력 2024.07.31. 20:02:25

[프레시안]

지난해 9월 국내 최대 포털사이트 네이버가 출시한 생성형 인공지능(AI) 검색 서비스 '큐:'(Cue:)가 '여성혐오', '페미니즘 사상검증' 등 일부 여성 의제 관련 키워드에 대해 '편향된 정보'라며 답변을 거부하고 있는 것으로 확인됐다. 사회적으로 민감한 의제인 만큼 답변을 거부하는 것이 소비자들이 받아들이는 데 가장 적합한 방식이라고 판단했다는 게 네이버 측 입장이다.

(덧붙여) 악성 댓글을 숨기는 기능의 '네이버 AI 클린봇'이 2020년 도쿄올림픽 중계 댓글에서 인종차별적 표현을 걸러내지 못하자, 네이버는 해당 표현을 비하로 보아야 할지 사람마다 기준이 다르며, 표현의 자유가 지켜져야 한다는 해명을 하기도 했음

차별·혐오의 사회적 기준 마련 (차별금지법 제정)

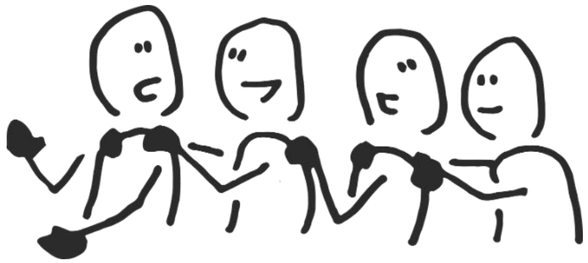


AI가 차별·혐오표현을 식별하여 걸러낼 수 있는 시스템 구축하고, 시스템이 제대로 작동할 수 있도록 차별·혐오의 사회적 기준을 세운다.

AI가 일으키는 차별·혐오를 방지하기 위해서는 무엇이 차별·혐오인지 AI가 식별할 수 있어야 함. 사회적 맥락에 따른 차별·혐오표현의 데이터셋을 구축하고, 차별·혐오표현을 걸러내도록 하는 등의 시스템이 필요. 그러나 AI가 모든 차별·혐오 표현을 식별하여 차단하는 것은 불가능하며, 그것이 차별·혐오표현에 대한 완전한 대책이 될 수도 없음. 따라서 차별·혐오표현의 사회적 기준을 세우는 일이 우선되어야 함.

질문하는 사람들 더 많이 조직하기

AI가 일으키는 차별을 심각하게 인지하고 이에 맞서고자 하는 사람들을 조직하자

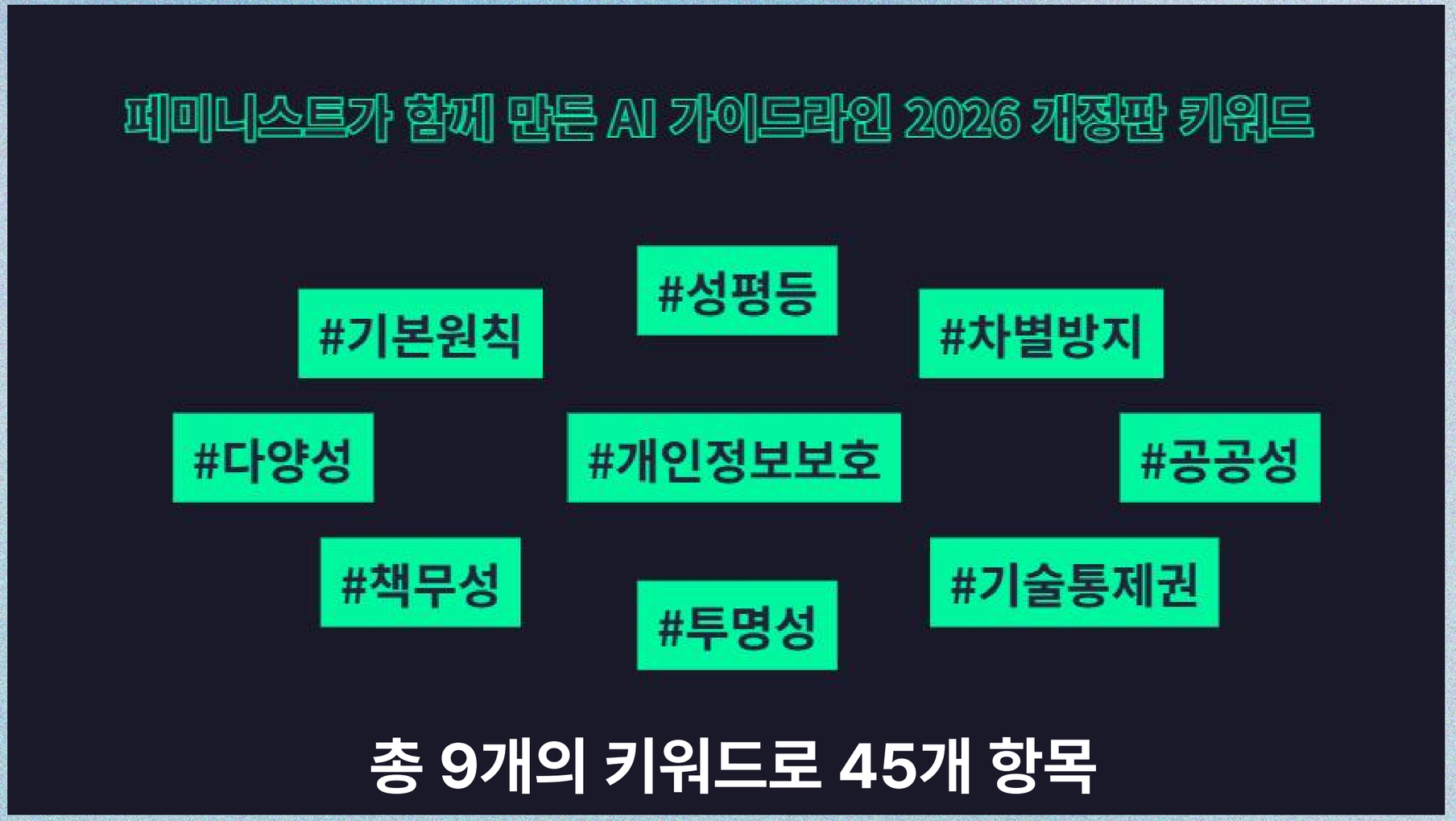


연일 AI의 편리함과 유용함에 대한 상찬이 이루어지고, AI 사회로의 선제적이고 전면적인 전환이 필연적이라는 담론이 널리 받아들여지고 있음. 이런 상황에서 잠시 멈추고 질문을 던져보자는 제안을 하기는 무척 어려운 일이기도 함. 하물며 그것이 AI가 일으키는 (성)차별을 비롯한 사회적 문제에 관한 불편한 질문이라면 더 그러할 것임. 그러나 바로 그 점이 우리가 더 질문을 던져져야 하는 이유이기도 함. '국민' 모두가 AI를 사용할 수 있다고 미디어 공공성이 확보되고 'K-민주주의'가 완성되는 것은 아닐 것임. AI가 이토록 전 영역에 필요한게 맞는지, 진짜 불가피한 것인지 계속 질문해야 함.

페미니스트가 함께 만든 AI가이드라인

Womenlink
한국여성민우회





자율살상무기, 누가 방아쇠를 당겼는가

AI시민사회포럼

패널 토론세션 트랙2 (책임성)

팔레스타인평화연대 자아

01
집단학살과 규범의 시차

- **‘가스펠’** : 감시 데이터(위성·드론 영상, 통신 감청 등)를 기계학습으로 분석. 2023년 10월 이전까지 이스라엘점령군이 가자에서 찾아내던 표적 : **1년**에 50개 → **하루에** 100개. 2023년 12월 기준 **하루 최대 250개** 타격
- **‘라벤더’** : 사람 단위로 통신·이동패턴·소셜 연결망 등을 종합해 "하마스/이슬라믹지하드 연계 가능성 점수"를 매겨 개인을 표적 명단에 올림. 최대 37,000명의 팔레스타인인 표적 명단에 올림. 한 명당 심사 시간 20초

- **CCW 자율살상무기(LAWS) 정부전문가그룹, 2026.9, 3년 논의 끝에 최종보고서 채택**
- 생산·비축·이전 제한, 효과를 예측할 수 없는 무기의 사용 금지, 개조 후 재심사, 교전 횟수 제안, 인간 승인 없는 표적 변경 금지 → **모두 삭제. 심지어 핵심 보호장치 삭제는 시민사회 접근이 차단된 비공식 협의에서 이루어짐**
- 사고 조사 의무 (자율무기체계가 오작동하거나 민간인 피해를 낸 사건이 발생했을 때, 국가가 그걸 조사하고 책임 소재를 가릴 국내 절차를 두게 하는 조항) → **‘보장’에서 ‘장려’로 후퇴**
- 규범 형성의 속도가 무기화와 학살의 속도를 따라가지 못하는 현실

출처: 피스모모, 「CCW GGE LAWS 2026년 2차 회의 결과 분석 보고서」(2026.10.19)

02
팔란티어 : 공통의 인프라



- **팔란티어 CEO 알렉스 카프** *“팔란티어가 우크라이나 전장 표적생성 (targeting)*의 대부분을 담당하고 있다. *Palantir is ‘responsible for most of the targeting in Ukraine’”* (Reuters, 2023.2.2)
- **미국의 이란 침공:** 팔란티어 메이븐 스마트 시스템 + 클로드, 24시간 내 1,000여 개 표적 생성
- **한국** : 조선·방산 두 축이 같은 인프라에 연결되어 있음
- **HD현대**: 2021년부터 팔란티어 Foundry 도입, 2024년 무인수상정 공동개발(정찰→전투용 확대 계획), 2026.1 수억 달러 규모 계약으로 그룹 전반 확대
- **LIG넥스원**: 2026.3 팔란티어와 UAE향 통합방공망·무인체계 MOU 체결
- **‘소버린 AI’ 관련 국내 정책 실정** : 한국이 이미 강한 메모리 반도체만 외국 기술 의존 없이 자국이 통제하는 ‘소유’ 전략
- 국방 C2·데이터 통합은 팔란티어, 앤듀릴 등 명시하며 ‘협력·접근’ 으로 분류- 현역 해군 대령/정책 설계자 스스로 인정하는 현실

Bae, Hackyoung. ‘Sovereign AI and ROK-U.S. Cooperation.’ Asia Policy 21.2 (2026): 107–124.

03
앤트로픽이 보여준 윤리 코스프레

- 2026.2 펜타곤 (이란 침공, 베네수엘라 침공 도중) 앤트로픽에 자율살상무기·대중감시 제한 해제 요구 → 거부 → 국방물자생산법(DPA) 발동하거나 '공급망 리스크' 지정하겠다 통첩, 6개월 내 국방부 시스템에서 퇴출 명령, 블랙리스트 지정
- 그러나 2024.11부터 이미 앤트로픽은 팔란티어의 국방 인프라 위에서 정보·국방기관에 Claude를 공급 중
- 2026.4 블랙리스트 기간에도 NSA가 미공개 모델 '미토스'를 중국·이란 대상 공격적 사이버작전에 사용하고 있다는 것이 보도됨. 엔지니어 6명까지 NSA에 상주 지원
- 2026.8 연방법원 "펜타곤이 블랙리스트 지정한 건 위법" 판결
- 2026.10 연방·주·지방정부 전체 대상 계약 오히려 확대
- 애초 세운 '레드라인'도 "미국 시민에 대한 대중감시"(즉 중국,이란은 감시해도 됨)", "완전자율무기(인간이 최종 결정에 전혀 개입하지 않는 것)"을 금지하겠다는 것으로 매우 좁게 설계 (인간이 최종 확인 누르는 라벤더식 표적생성은 처음부터 금지 대상이 아니었음)

04
방아쇠를 당긴 것은 누구인가

- 팔란티어 같은 방산 AI를 키우고 앞세우는 미국
- '원칙'은 자기 사업 키우는 협상 카드로만 활용하는 앤트로픽 등 플랫폼
- 그 구조에 자발적으로 올라탄 한국의 행위자 - HD현대·LIG넥스원·‘소버린 AI’ 정책
- 한국 시민사회는 무기 수출만이 아니라 전쟁 수행 인프라(AI, 데이터, 병력) 전반에 대한 가담을 거부해야함

AI 대전환 시대,
AI는 성평등 할까?

[AI시민사회포럼] 폭주하는 AI를 멈춰라 - AI의 공공성과 책임성을 묻다(2026.11.6), 트랙 2 책임성 부분 발표자료

임선희(한국여성단체연합 사무처장)

“

한국 정부
AI 3대 강국 발표.
규제 완화와 산업 육성
이면에는 무엇이 존재할까?

AI가 야기하는 차별과 폭력

1 데이터와 알고리즘의 젠더 편향

학습 과정에서 사회에 만연한 성차별과 성별 고정관념도 함께 학습, 이로 인한 혐오와 차별 재생산

2 AI 기술을 이용한 젠더폭력

딥페이크 기술 등을 활용한 성착취물 대량 생산, AI 봇 등을 통한 성매매 조장/홍보 등 범위 확대, 피지컬 AI의 섹스봇화

3 과잉 정보 집적, 개인정보 침해

동의없는 개인정보의 과도한 수집, AI에 의한 의료기관, 학교, 은행 등의 광범위한 해킹

4 고용 불안정과 노동환경의 변화

AI 기술 적용이 노동, 복지 등 광범위하게 확산

AI의 성편형과 차별

인공지능 비서 기본 설정값은 여성

- 예) 애플 시리, 아마존 알렉사, 구글 어시스턴트
- 기본 설정값은 여성의 목소리
- 유네스코* '여성화된 AI 비서는 성별 고정관념을 강화함'
- 애플 : (2013년) 여성의 목소리를 기본값으로 설정하지 않고 사용자가 선택하도록 함. (2021년) 2개 목소리 추가

AI 채용 시스템의 성편향

- 아마존 AI 채용시스템 2014년 도입
- AI 채용 시스템이 이력서에 '여성'이라는 단어가 포함되면 감점
- 특정 용어에 불이익을 주지 못하도록 개선했지만, 같은 문제가 반복되지 않으리라는 확신을 주지 못해 폐기(2018년)

AI 챗봇의 소수자 차별/혐오

- AI 챗봇 이루다 (2020년 12월)
- 여성 차별, 장애인, 성소수자, 흑인 혐오 발언으로 오픈 3주 만에 서비스 중단
- 서비스 개발 시 학습과정에서 카카오톡 대화 내 이름, 휴대전화번호, 주소 등 개인정보에 대한 삭제/암호화 등 조치 전무. 약 60만 명 이용자의 대화문장 94억여 건 이용

*UNESCO, I'd blush if I could: closing gender divides in digital skills through education, 2019

AI 기술에 대한 우려와 문제점

- [라운드테이블] A.I.와 성평등 - A.I.가 가져올 사회 변화, 젠더로 묻다(한국여성단체연합, 2026.05.06. 전체토론 주요 내용)

젠더폭력

- 성폭력 사건해결 과정에서 생성형 AI의 사용자 친화적인 답변 행태가 성폭력 피해자에게 거짓 정보 제공 또는 잘못된 판단을 이끌어낼 우려
- 가해자가 AI로 작성한 반성문 등

여성 노동

- 서비스, 돌봄, 사무, 상담 등 여성이 집중되어 있는 일자리(콜센터)가 AI로 대체
- 여성노동자들이 수행해온 노동의 숙련성이 인정되지 못하는 상황
- 노동자 감시 문제

섹스로봇

- 성인형 인형 리얼돌과 섹스로봇 개발 회사 헬스케어 기업과 합병(2026년)하며 '섹스'를 지우고 'AI 휴머노이드'라는 용어를 내세우며 소위 '가전제품'이나 '첨단 수출품'으로 취급하여 규제 회피
- 섹스로봇과 사용자의 관계가 공감을 전제하지 않고 일방적으로 이루어지는 것은 성매매와 유사한 구조

현재 인공지능기본법의 문제점

성평등 가치, 차별/폭력 금지에 대한 원칙 필요

차별과 혐오, 폭력 방지에 대한 사회적 책무 규정 필요

- 기본원칙 및 국가의 책무(제3조), 인공지능 윤리원칙(제27조)에 추가

의사결정과정에서의 여성의 참여, 성별균형 보장 필요

- 위원회 성비 원칙을 양성평등기본법에 부합하도록 수정
- 시행령 제4조에 성평등가족부 추가

※ 참고 : 인공지능기본법
제7조(국가인공지능전략위원회) ② 개정(특성 성이 10분 6 초과하지 못하도록)
제28조(민간자율인공지능윤리위원회의 설치 등) ③ “민간자율위원회의 구성·운영 등에 필요한 사항은 해당 기관 또는 단체 등에서 자율적으로 정한다. 다만, 그 구성에서 특정한 성이 10분의 6 이상이 될 수 없으며”로 수정

감사합니다.

AI가 일감을 정할 때, 누가 책임지는가?

플랫폼의 AI·알고리즘은 단순한 기술이 아니라
노동자의 기회와 소득에 영향을 미치는 결정 체계입니다.

핵심 질문 | 그 결정을 노동자가 이해하고, 다투고, 바꿀 수 있는가?

알고리즘과 AI: ‘계산’에서 ‘결정’으로

알고리즘

규칙에 따라 배차·순위·정산 등

결정을 실행하는 절차

AI 결합

데이터로 수요·시간·행동을 예측해

배분과 평가를 더 자주 조정

중요한 것은 AI 사용 여부보다 노동조건에 미치는 실제 영향

노동자의 데이터가 다음 노동조건을 결정한다.

일하기

수행·이동·대기

→

기록하기

시간·위치·평가

→

예측하기

수요·성과·위험

→

결정하기

배차·보수·제재

기업은 전체 데이터와 기준을 보지만 노동자는 자신의 결과만 보는 경우가 많습니다.

배달·대리·택배: 기술보다 '결정의 결과'가 중요하다.

배달라이더

배차·예상시간·묶음배송

→ 수입과 이동속도에 영향

대리운전

호출 노출·배차·평가

→ 일감 접근과 대기시간에 영향

택배

물량 예측·경로·실적 관리

→ 작업량과 현장 자율성에 영향

※ 기업별 기술 적용 범위와 영향은 공개자료·현장자료로 별도 확인해야 합니다.

같은 알고리즘, 기업과 노동자의 질문은 다르다.

기업의 질문

어떻게 더 빨리 배달할까?

어떻게 수요와 공급을 맞출까?

어떻게 비용과 위험을 줄일까?

노동자의 질문

왜 나는 이 일을 배정받지 못했나?

왜 보수가 달라졌나?

불리한 평가를 누가 바로잡나?

책임성 = 효율성뿐 아니라 결정의 정당성·설명 가능성·구제 가능성

문제는 ‘자동화’보다 ‘책임 없는 자동화’다.

① 불투명성

어떤 기준으로
불이익이 발생했는지
알기 어렵다.

② 이의제기 한계

오류·차별·오탐에
사람이 책임 있게
답하지 않을 수 있다.

③ 책임 분산

플랫폼·위탁업체·
자동화 시스템 사이에
책임이 흐려진다.

AI가 판단을 돕더라도 그 결과에 대한 책임은 기업과 제도에 남아야 합니다.

AI 시대 노동권: 네 가지 최소 조건

알 권리

결정 목적·핵심 기준·변경 고지

설명받을 권리

불이익 결정의 이해 가능한 이유

다툴 권리

사람의 재검토·정정·구제

참여할 권리

노동자 대표의 사전 협의·검증

국제 기준도 자동화된 노동결정의 책임을 묻는다.

EU 플랫폼 노동지침

자동화된 모니터링·의사결정에
대한 정보 제공, 인간 감독,
검토·설명 등 규율

한국의 정책 과제

플랫폼 종사자의 법적 지위와
무관하게 중요한 노동결정의
설명·이의제기·참여 보장

참고: Directive (EU) 2024/2831 · 조문별 적용 범위 및 국내 이행은 구분 필요

책임을 누구에게, 어떻게 물을 것인가?

플랫폼 기업

설계 목적과 영향 평가

중대한 결정의 설명

인간에 의한 재검토

정부·감독기관

권리의 최소기준 마련

자료 제출·감독·시정

분쟁 구제 절차

노동자·대표조직

현장 영향 공동 검증

변경 전 협의·교섭

집단적 대응과 지원

AI 책임성의 기준을 노동현장에 적용하자 !

1. 중대한 배차·보수·제재 결정의 기준 고지
2. 자동화된 불이익에 대한 인간 재검토
3. 안전·소득·차별 영향의 정기적 점검
4. 노동자 대표와의 사전 협의·집단 검증

AI의 책임성은 노동자의 권리에서 검증됩니다.

노동자의 일감을 결정할 권한이 있다면,
그 결정의 이유를 설명하고 결과를 바로잡을 책임도 있어야 합니다.

알고리즘의 정확성만이 아니라, 노동자의 존엄·안전·참여권이 기준이 되어야 합니다.

복지알고리즘의 차별과 배제를 어떻게 막을 것인가

*본 자료는 <사회복지분야 인공지능 도입현황과 과제　영향받는 사람들 연속 워크숍>(2026.9.2) 빈곤사회연대, 홈리스행동, 전국장애인차별철폐연대, 디지털정의네트워크, 공익인권법재단 공감 등이 함께 꾸린 자료를 발췌해 작성되었습니다.

- 빈곤사회연대 김윤영

사회보장분야 ai도입 키워드

#사각지대(발굴)

- 사각지대 발굴 위해 21개 기관 47종 위기정보 연계, 단전, 단가스부터 각종 복지제도 이용과 체납, 실직과 휴폐업여부, 병원이용 기록까지 매월에서 격월로 정보 수합
- 발굴대상자는 각 지자체로 수합되고, 일선 담당자는 현황을 파악, 지원연계와 결과 리포트
- 발굴된 고위험가구 중 단기 민간서비스가 아닌 기초보장, 차상위, 긴급복지로 연계된 경우는 11.1%
- 2024년 11월부터 ai를 활용한 초기상담 진행 중.

#자격관리

- 2010년 사회보장통합전산망 도입. 7개 부처 170개 복지서비스 이력과 국세청, 노동부, 건보공단 등에 등록된 소득재산자료 및 인정정보가 모두 담김. 복지수급자의 소득재산 변동사항 역시 통합관리.
- 도입 이후 2010년과 2011년 복지급여 대상자 확인조사 실시, 44만 8천 9백명의 급여 박탈. 이 중 기초생활 수급자는 11만 7천명.
- 2026년 11월부터 근로사업소득 조사 ai로 진행. 복지행정업무 50%이상 자동화를 목표로.

#돌봄

- '복지, 돌봄 ai대전환 로드맵' 마련 중. 생활반응이 없거나 낙상사고 등을 감지하는 '스마트 홈' 구축. 이를 위한 R&D투자 확대에 방점.
- NHN 와플랫 AI 생활지원사. 전국화를 목표로 각 지자체 및 공공분야과 협업 중. 스마트폰 기반으로 24시간 비대면 안부확인과 생활관리 지원하는 ai통합돌봄 서비스. 안부 및 건강데이터를 상시 수집, 분석함.
- 페르소나는 50대 여성 생활지원사.

빈곤 현장에서 디지털 전환이 낳는 쟁점들

1) 오류로 인한 피해와 수급자의 입증책임 강화. 디지털 처리의 복잡성과 일방성은 수급자의 이의신청이나 구명활동을 어렵게 만들고, 공적 자료와 실제 삶 사이의 차이를 해명할 책임이 수급자에게 돌아옴.

2) 제도 밖의 예외 사례를 구제할 동기의 상실. 공적 자료를 중심으로 적격 여부를 판단하는 것으로 전담공무원과 일선 사회복지사의 역할이 제한되고, 적극적으로 제도의 목표(사회적 권리의 보장)를 실현할 필요가 약해짐.

3) 과도한 정보공개와 정보노출을 피하려는 사람들을 복지제도 이용에서 배제. 사회적 권리를 위해 데이터를 보호받을 권리를 포기해야하거나, 사생활을 방어하려 할 때 복지대상으로 '발굴되지 않는 책임' 또한 수급자의 몫이 됨.

4) 부정수급, 급여 삭감을 중심으로 한 제도 운영 강화. 일제조사와 같은 대량 감시가 잦아지고, 이는 수급자들로 하여금 급여를 받는 상태가 적정하지 않다는 인식을 강제함.

5) 정보유출 가능성. 제동한 정보가 사회복지급여 책정 외 다른 방식으로 활용될 수 있고, 특히 유출되거나 판매될 경우 저소득층에게 저질의 상품, 고리의 대출이 찾아가는 등 심각한 피해가 발생.

먼저 온 위험의 징후

#장애

- 서울시 등은 발달장애인평생교육센터에 비전AI기반 CCTV를 설치, '도전행동'을 자동으로 포착, 기록.
- 이렇게 기록된 정보는 알고리즘에 의해 '위험할 수도 있는 문제행동' 또는 '문제행동 직전의 행동'으로 분류됨.
- 이는 얼마나 위험한 행동으로 분류하느냐에 따라 권리를 제한하는 것으로 대안이 제안될 수 있음.

#홈리스

- 전체 시설노숙인 중 등록장애인은 50.2%인데 장애유형 중 지적장애(20.0%)와 정신장애(19.0%) 비중이 타 유형에 비해 높고, 26.6%가 초졸 이하, 47%가 중졸 이하. 비대면 복지는 이들에게 중첩된 취약성을 발생시킴.
- 코로나19 팬데믹 당시 서울시는 '방역 관리'를 목적으로 무료 급식장에 전자태그(RFID)방식 회원증 도입. 본인의 기록을 노숙시설에 남기지 않으려는 이들은 급식도 제한되는 효과.
- 캘리포니아에서는 공공장소에 주차된 노숙 차량을 포착하는 AI기술 개발하기도.

#청소년

- 성평등가족부 산하 한국청소년상담복지개발원에서 '청소년상담1388' 운영 과정에서 수집한 상담기록 1만 4천여건을 민간업체에 제공.(2026.10)
- 여기에는 청소년들의 임신, 성매매, 자살, 자해, 가출, 가정폭력, 성폭력, 학교폭력 등 상담 기록을 포함. 사전 개인동의나 사후 제공 사실 고지 없이 인공지능 학습용으로 기업에 제공.
- 디지털헬스케어법은 개인동의없이 가명처리만 하면 의료, 건강정보를 기업에 넘기도록 하고 있음.

소결: '기술'이 아니라 '정치적 의지'가 권리를 지킨다

- 대규모 감시와 정보 유출도 문제지만, 이에 대한 보완책으로 '인간의 최종적 결정과 책임'을 남겨두면 된다는 관점 역시 안이할 수 있음. 과도한 정보공개와 노출을 피하려는 이들을 사회복지에서 배제하고, 부정수급이나 급여삭감을 중심으로 한 대량의 감시와 통제가 용이하게 만들기 때문.
- 이는 개인에게 더 많은 책임을 떠넘기는데, 더 나은 디바이스와 구독서비스 구매로 스스로 격차를 해소하라거나 복지로부터 지원받아야 할 필요성을 스스로 입증하라는 것은 동일한 원칙 아래 작동함. 제도에 의해 권리가 충분히 보장되지 않는 상황에서 '지원받을 자격이 있음'을 입증하는 것은 필연적으로 낙인과 실패를 동반함.
- 즉, 복지 사각지대나 권리의 부족 상태에 대한 기술주의적 해법은 문제를 해결하지 않을 뿐만 아니라 '기술로 권리를 대체할 수 있다'는 착시를 가져옴.
- 그렇다면 기술이 놓여야 할 자리는 어디인가? 우선 '기술은 기술대로, 제도는 제도대로 논의해야' 한다거나 '기술이 더 발달하면 제도의 문제 해결할 것'이라는 식의 낙관부터 버려야 함. 제도의 목표에 관한 정치 사회적 합의와 이를 통한 권리보장을 우회할 수 없음.

의료 AI의 오진과 구조적 책임

이서영(건강권실현을위한보건의료단체연합 정책위원)

의료 AI 오진을 둘러싼 담론과 구조적 책임

- 주류 '책임성' 담론의 주제들 : 의료인 과실책임 여부, AI 의료기기 제조물책임, 의료기관의 사용자책임 등
  - 의료계의 대응: AI 활용 과정에서 의료인/의료기관의 법적 책임을 회피·최소화하기 위한 논의, 인공지능 자체의 다층적 의료윤리적 문제를 다루지 못하고 있음
- 문제의식 : 의료 AI 도입 과정을 자본의 이해관계에 따른 의료노동 재편의 과정으로 이해한다면, 의료인 개인의 주의의무를 넘어 기술의 생산·도입·운영에 관여하는 자본과 국가의 책임을 논의하는 것이 핵심이 되어야 함
  - 핵심 쟁점: 건강불평등을 재생산하는 의료 AI를 둘러싼 사회적·경제적 구조에 대한 책임은 누구에게 있는가?
- 이 발제에서는 정부의 무책임한 의료 AI 규제완화와의료 현실의 부조화에 대해 집중적으로 제시해보고자 함.

예견된 오진 : 의료 인공지능의 기술적 현황

- 생성형 AI는 현재까지 아직 진단정확도와 성능 자체에서 기술적 한계를 보이고 있음
- AI의 편향은 기존 의료체계의 불평등을 학습하고 재생산할 수 있음.
  (누락된 의료기록, 의료인의 진단 편향, 사회경제적 지위에 따른 의료서비스 격차가 AI의 학습 데이터에 반영됨)
- 대규모 언어 모델(LLM)은 유해한 의학적 허위 정보를 흡수함. 특히 권위 있는 임상 문체로 표현된 경우 더욱 그러한 경향을 보임.
- 오진이 대규모로 발생할 수 밖에 없는 기술을 현장에 도입한 것은 누구인가?
  이것으로부터 이윤을 거두는 집단은 누구인가?
  대규모 오진 가능성을 유발하는, 책임성을 가진 대상은 누구인가?

정부의 책임 : 무분별한 규제완화

- 정부는 적극적으로 산업을 촉진하는 역할을 자처, 의료기관들에 인공지능 도입을 촉구, 규제완화에 나서고 있음.
  - 신의료기술 평가유예: 비침습적 AI 진단 소프트웨어 등의 평가를 유예하고 최대 3년간 현장 사용 허가
  - 시장 즉시진입 제도: AI 진단 보조 기술 등 특정 품목군에 대해 별도 평가 없이 최소한의 확인만으로 3년간 즉시 시장에 진입할 수 있도록함.
- 이렇게 의료현장에 도입된 인공지능은 효과성을 입증하지 못했음에도 불구하고 환자에게 많은 비용을 부과하고 있음. ([단독-빗장 풀린 의료AI ①] 검증 끝나지 않은 의료 AI, 비용은 환자 몫 - 정식 신의료기술평가 통과 0건 … 공급가보다 최대 2.6배 청구, 매일노동뉴스, 2026-10-13)
- 그런데, 이상사례 보고는 0건? : 시장 도입 이후 환자안전과 의료환경에 미치고 있는 영향평가는 제대로 되고 있지 않음 (AI 의료기기 안전관리 사각지대 우려 - 수백 건 허가에도 이상사례 '0건'…오진·판독 누락 등 개선책 필요, 의학신문, 2026-10-12)

지역의료 공백의 현실

연합뉴스

보은지역 유일 병원 응급실 내달 1일 폐쇄…의료공백 불가피



(보은=연합뉴스) 박병기 기자 = 충북 보은지역의 유일한 보은한양병원 응
내달 1일 문 닫을 예정이어서 응급환자 관리에...

2일 전

남도일보

전남권 산부인과 진료 공백 '빨간불'

전남에서 산부인과 전문의가 없는 지역이 현재 곡성군 1곳에서 10년 뒤
곳으로 늘어날 수 있다는 분석이 나왔다. 의료인력...

27분 전

인천투데이

강화·옹진 의료공백…섬 이송 최장 9시간 "2028년
항"

허종식, 27일 국감서 인천 섬·접경지역 의료 공백 지적 강화 공보의 정원
무 만료 섬 응급이송 시간 전국 2.4배… 닥터...

4일 전

의협신문

지역·필수의료 공백 메울 시니어 의사, 예산에 발목

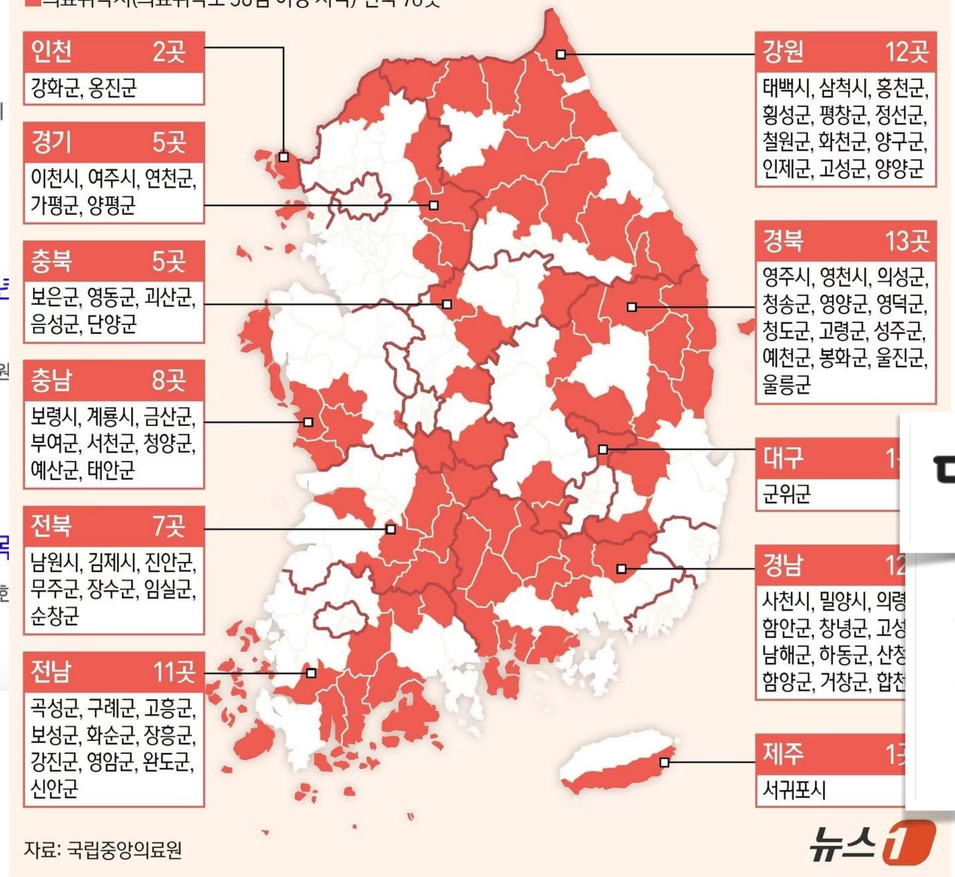
지역·필수의료 부족 문제 해결책으로 꼽히는 '시니어 의사제도'가 예산 한
러났다.더불어민주당 소병훈 의원은 보건복지부...

1주 전

2025년 기준 전국 의료취약지

※분석 지표: 종합병원 60분 접근불가능 인구비율, 종합병원 60분 기준시간 내 의료이용률)

■의료취약지(의료취약도 50점 이상 지역) 전국 76곳



자료: 국립중앙의료원

뉴스1



매일노동뉴스

사람 줄이고 AI 늘리고, 거꾸로 가는 복지부 예산

실집행률 8.8% AI에는 예산 4배 증액, 인력은 동결

우다영 기자 입력 2026.10.30 06:30

사회운동의 당면과제와 주장

- 안전과 효과가 입증되지 않은 의료 AI 기술 도입 중단
- 사회적 차별과 편견을 강화할 수 있는 AI 기술에 대한 통제
- 미입증된 기업 이윤추구의 수단인 AI 기술 도입보다, 확실하게 입증된 공공의료 확충, 보건의료 인력 확충
- 차별과 편견을 강화하고 환자를 위험하게 만들 수 있는 의료 AI 기술을 개발한다는 명목으로, 개인 건강정보까지 기업에 넘겨주려는 '디지털헬스케어법' 중단

참고문헌

- Takita, H., Kabata, D., Walston, S.L. et al. **A systematic review and meta-analysis of diagnostic performance comparison between generative AI and physicians.** npj Digit. Med. 8, 175 (2025). https://doi.org/10.1038/s41746-025-01543-z

- Cross JL, Choma MA, Onofrey JA. **Bias in medical AI: Implications for clinical decision-making.** PLOS Digit Health. 2024 Nov 7;3(11):e0000651. doi: 10.1371/journal.pdig.0000651. PMID: 39509461; PMCID: PMC11542778.

- Omar M, Sorin V, Wieler L et al. **Mapping the susceptibility of large language models to medical misinformation across clinical notes and social media: a cross-sectional benchmarking analysis**, The Lancet Digital Health, 8

‘환자 비밀’을 팔아 돈을 벌겠다는
‘디지털 헬스케어법’ 반대!

여러분,

지금 국회에 우리가 아파서 병의원 약국에 남긴
진료기록과 처방내역, 검사결과, 복약정보 등이 우리 동의도 받지 않고,
정부 맘대로 기업에 제공하는 법안이 올라와 있습니다.
이름하여, **‘디지털 헬스케어법’**입니다.

디지털헬스케어 산업을 발전시키고, AI의료를 발전시킨다는 명목으로,
‘환자의 민감한 진료기록과 건강정보 보호를 해제하는 법안’입니다.

한번 유출된 의료·건강정보는 비밀번호처럼 바꿀 수 있는 게 아닙니다.
가리고, 숨기고 싶어도 다시 숨길 수 없으며 지울 수도 없습니다.
평생을 쫓아다니며 낙인과 차별, 그리고 불평등을 만들 수 있습니다.

우리의 소중한 의료 기록과 건강정보를 지키기 위해

함께 합시다!



SCAN ME

법안 폐기 요구 서명에 함께해요!

추천 알고리즘과 중독적 설계

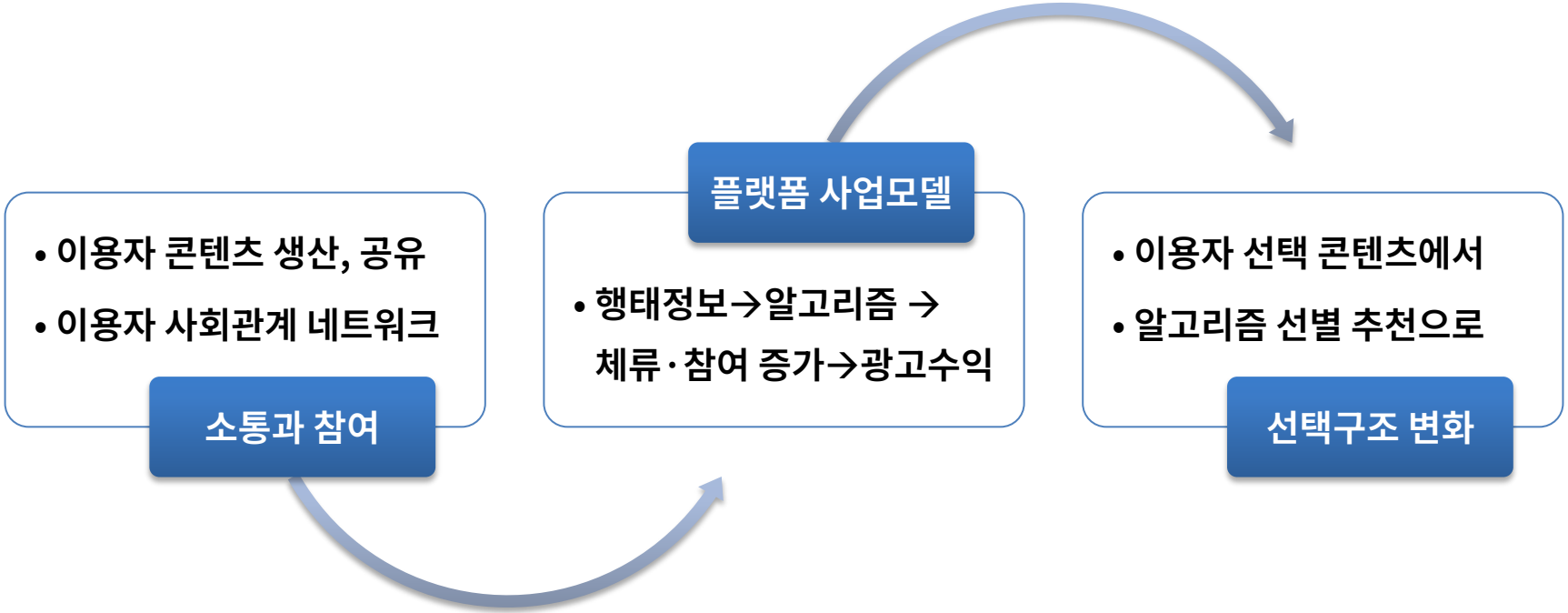
2026. 11.

장여경

사단법인 정보인권연구소 Institute for Digital Rights

소셜미디어 플랫폼

○ 이용자 소통과 참여의 공간이자 이용자 주의력과 데이터를 수익화하는 플랫폼



* 자료: L. Wiesböck et al. (2026). ADDICT: Addictive Design in Digital Consumer Technologies. IHS(Institute for Advanced Studies) & AK Wien, Materialien zur Konsumforschung, VOLUME 20.; C. Entrena-Serrano. (2025). Watch, Scroll, Repeat: How Interface Design Shapes Consumptive Curation Affordances on TikTok. Social Media + Society, July-September 2025: 1–16.

알고리즘으로 추론한 정체성: '추론된 나'



추천 알고리즘과 중독적 설계

○ 무엇을 보여줄지 결정 + 계속 이용하도록 유도하는 설계

- 추천 알고리즘: 시청시간·클릭·반응 등 행태정보를 분석하여 콘텐츠의 노출과 순위를 결정
- 멈춤 신호 제거: 무한 스크롤·자동 재생으로 이용을 중단하고 숙고할 기회를 축소
- 반복 접속 유도: 푸시 알림·새로고침·좋아요·사회적 인정과 FOMO를 결합

➔ 중독적 설계

- 강제적 행동 · 영구적 사용성 · 게임화 · 설계된 사회적 연결 · (초)개인화 · 시간적 구조화 · 인플루언서 수익화

사회 정체성의 게임화와 민주주의 영향

- 소셜미디어 환경은 일반화 단계를 생략하고 사회적 지위를 추구하는 정체성 형성
  - 정치적 판단의 일반화 단계(아렌트): 자신과 다른 관점을 일반적으로 상상하고 자기 판단의 정당성을 검증하는 비판적 단계
  - 반면 정치적 정체성에 기반한 사회적 지위를 추구하면 일반화 단계를 생략하고 극단화
- 소셜미디어는 유사 정체성 집단으로부터 정량적 피드백/[여론]입증으로 보상
  - 페이스북 좋아요, 게시판 좋아요/공감지수 등
  - 사회 정체성의 '게임화 ' ⇒ 동질집단 비교에 헌신 ⇒ 극단화

* 자료: Daus, Z. (2024). Socializing the political: rethinking filter bubbles and social media with Hannah Arendt. Ethics Inf Technol 26, 20. https://doi.org/10.1007/s10676-024-09759-5

취약성 위험

- 개별 유해 콘텐츠나 이용자 개인의 노력 여부에 따른 부작용을 넘어 기업이 행동을 유도하는 설계 구조의 문제
- 이용자의 약한 지점을 찾아내 참여와 수익으로 바꾸는 설계: 취약성을 표적화하거나 취약성에 영향을 미치는 체계적인 위험
  - 연령적 취약성: 아동·청소년의 발달 특성뿐 아니라 고령자의 디지털 이해·통제 역량 차이도 고려
  - 심리적 취약성: 외로움·불안·사회적 인정 욕구·놓치는 데 대한 두려움(FOMO)을 이용해 반복 이용을 촉진
  - 경제적 취약성: 금전적 압박과 보상 기대를 이용한 긴급성·희소성 메시지, 플랫폼 노동 종속성 강화, 충동적 소비 유도

➔ 플랫폼 기업-개별 이용자 간 체계적인 권력 불균형을 시정하기 위해 공공 개입 필요

알고리즘 설계를 원인으로 하는 기업 책임

- 공공적 개입: 기업의 설계 결정과 수익구조를 제한
- 원칙적 제한
  - 무한 스크롤·자동 재생·야간 알림 등 중독적 기능 원칙적 제한, 참여도 기반 추천 제한-비개인화 피드 선택 등
- 예방적 책무
  - 취약성 보호와 안전한 기본 설정, 사전적 영향평가와 조치 의무 등
- 알고리즘 투명성과 통제권
  - 추천 기준 설명, 프로필 수정·추천 초기화, 손쉬운 중단 등
- 기업 책임 공공 감독
  - 독립적 감사, 영향받는 집단 참여, 시정 및 제재, 피해구제, 입증책임 전환 등

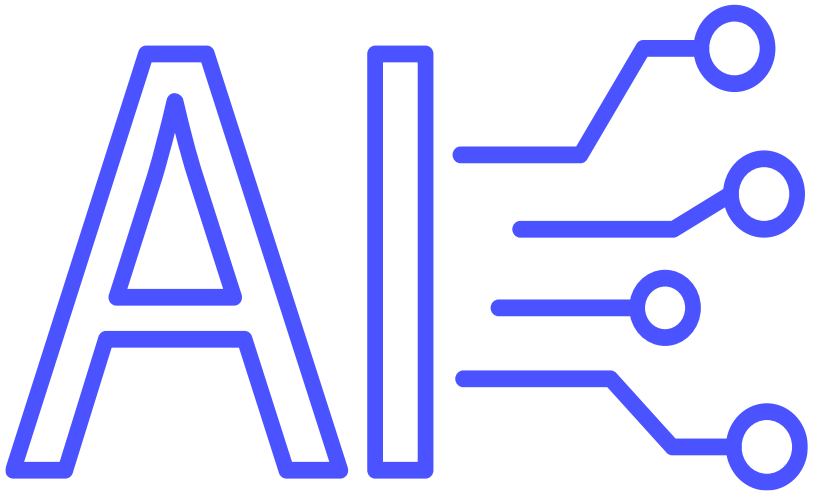
감사합니다

AI시민사회 포럼 트랙 2(책임성)

AI 위험성 식별 및 완화를 위한 AI 인권영향평가의 필요성

민주사회를 위한 변호사 모임 디지털정보위원회
변호사 김병욱

인공지능에 의한 침해의 특성 및 위험성



- 인공지능 기술에는 기본권 전반에 관하여 부정적인 영향을 미칠 가능성이 잠재되어 있음.
  - 개인정보의 무단 수집, 활용
  - 편향된 데이터에 의한 차별적인 결정
  - 노동력의 대체, 노동과정의 재편
  - 사람의 생명, 신체의 안전을 침해할 우려 등.
- 매개변수의 생성과정은 불투명하고, 자율성을 지님(블랙박스 알고리즘)
- 인공지능의 자율적 특성, 불투명성, 시스템 개발과 활용에 따른 영향, 파급효 등을 고려할 때 사후적인 피해구제나 제재만으로 가 어려운 특성을 지님.

=> 인공지능 기술이 도입되기 이전 단계에서부터 사전에 위험요인을 식별하고, 제거하기 위한 체계적인 시스템이 필요함.

인공지능에 대한 인권영향평가의 필요성

"(인공지능에 의한) 많은 추론과 예측은, 사람들의 자율성과 자신의 정체성에 대한 세부사항을 확립할 권리를 포함하여, 프라이버시권의 향유에 깊은 영향을 미친다. 이는 또한 사상과 의견의 자유에 대한 권리, 표현의 자유, 공정한 재판 관련 권리 등 다른 권리에도 많은 문제를 야기한다. "

"인공지능 시스템의 설계, 개발, 배치, 판매, 구입, 운영의 수명주기 전반에 걸쳐 체계적으로 인권실사를 수행한다. 그 인권 실사의 핵심 요소는 정례적이고 포괄적인 인권영향평가여야 한다."

- 유엔 인권최고대표 <디지털 시대 프라이버시권 (2021)> -



"39. 국가는 인공지능의 개발과 활용에 있어서 인권침해와 차별의 가능성 및 정도, 영향을 받는 당사자의 수, 사용된 데이터의 양 등을 고려하여 공공기관 및 민간기업을 대상으로 인권영향평가를 실시하여야 합니다."

국가인권위원회 <인공지능 개발과 활용에 관한 인권 가이드라인> (2022.4)

인공지능 인권영향평가 도입 현황(해외와 국내)

- 유럽연합은 위험기반 접근법에 따라 인공지능시스템을 수인불가한 위험, 고위험, 제한된 위험으로 차등화하고, 고위험 인공지능 시스템에 대해서는 기본권 영향평가 시행의무를 포함하여 강한 수준의 의무를 부과하고 있음(AI Act 27조).

- 이외에도 네덜란드의 기본권 알고리즘영향평가, 덴마크 인권영향평가 등 다수의 국가에서 이미 인공지능에 대한 영향평가 또는 인권영향평가를 시행하고 있거나 그 도입을 예정하고 있음.

- 그러나 현행 인공지능 기본법에서는 인공지능 영향평가를 노력의무로 규정하고 있음(인공지능 기본법 35조).

- 국내에서도 윤리기준, 자율점검표 등 인공지능에 특화된 자율 점검기준이 있으나, 점검항목의 단순 나열에 불과하거나 객관성을 담보할 수 있는 장치가 부재한 실정임.

인공지능 인권영향평가



- 【1단계 : 계획 및 준비】
- 【2단계 : 분석 및 평가】
- 【3단계 : 개선 및 구제】
- 【4단계 : 공개 및 점검】

<국가인권위원회 '인공지능 인권영향평가 도구>

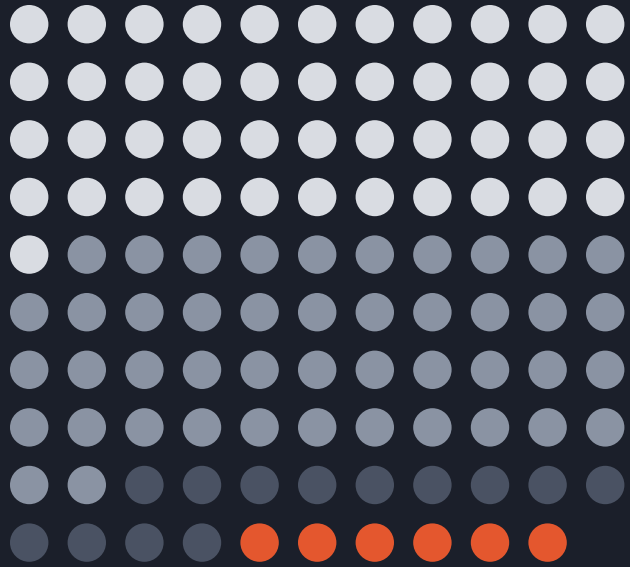
감사합니다

2026 AI시민사회포럼 발제 · 2026. 11. 6.

민주적 AI 거버넌스를 어떻게 구축할 것인가

이재근

참여연대 협동사무처장 · AI시민행동 공동집행위원장



국가AI전략위원회 민간·분과·특별위원 109명
밝은 회색 기업·산업계 · 회색 대학 · 주황 비영리·포럼

※ 위원 명단 분석·취합, PPT 작성에 AI(Claude)를 활용했습니다

‘AI 전성시대’, 시민사회의 자리는 어디에 있나

국가 정책의 중심이 된 AI

3대 메가프로젝트, 메가특구법 등 정부가 내세우는 정책 상당수가 AI와 연결되어 있다

일상과 정책결정까지 침투

대통령의 국무회의 ‘제미나이’ 인용 해프닝 — AI가 정책 결정 과정에 얼마나 들어와 있는지 보여준 사례

진짜 문제

AI 진흥·규제라는 거대한 전환 과정에 시민사회가 참여하거나 개입할 방법과 여지가 많지 않다

기대보다 공포가 큰 변화 앞에 서 있다

국가인공지능전략위원회: 대통령 직속 AI 정책 컨트롤타워

2025. 9
대통령령으로 출범

2026. 1
AI기본법 시행,
법정기구로 전환

2026. 2
인공지능 기본계획
(2026~2028) 의결

2026. 3
조직 개편
(분과·특위·TF 확대)

위원장	상근부위원장	간사	정부위원	분과·특위·TF
대통령	**민간 · 하정우 (실제 운영 담당)**	**대통령실 AI미래기획수석 (이해민)**	**대통령,과기부 장관 등 20명**	**10개 분과 2개 특위·1개 TF**

AI기본법 제7조: 인공지능 발전과 신뢰 기반 조성 등을 위한 주요 정책을 심의·의결

민간·분과위원 109명 중 82%가 기업·대학 출신

82%

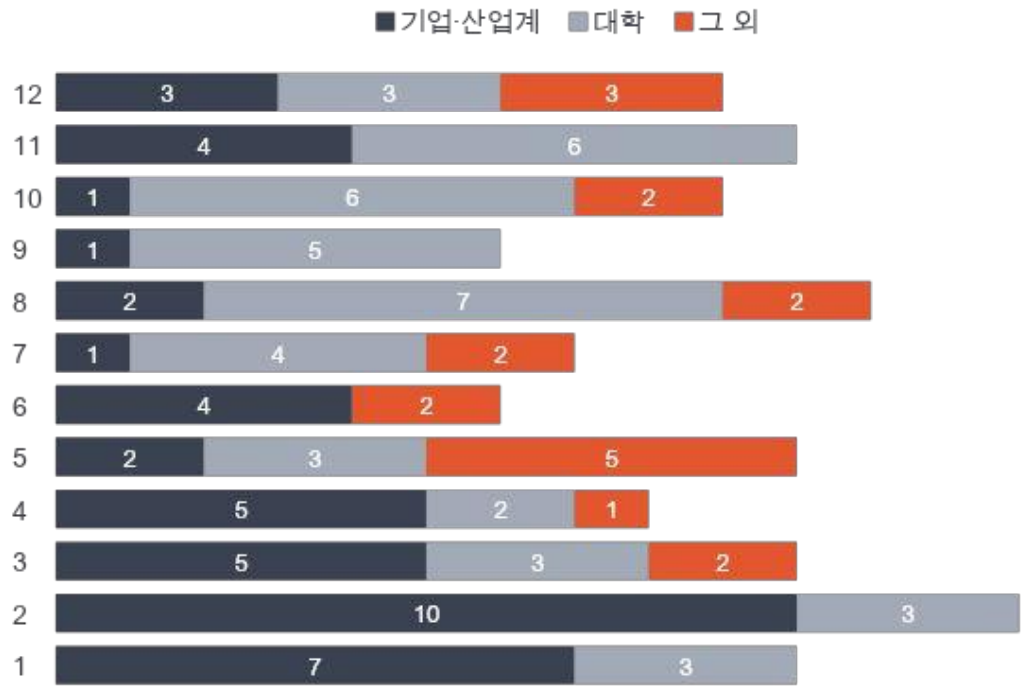
기업·산업계 45명 + 대학 45명
= 109명 중 90명

위원회는 3월 개편 때 127명이라 발표했지만 홈페이지에서 확인되는 위원은 109명



단위: 명 · 현직 대표 직함 기준, 중복 제외 · 위원회 홈페이지 '위원/분과위원/특별위원 소개' (2026.10.26. 확인)

분과·특위별로 보면 쏠림은 더 뚜렷하다



산업AX·생태계

13명 중 10명이 기업 출신

데이터 분과

개인정보보호 전문가로 볼 사람이 없다

보안 특위

10명 전원 기업·학계 출신

그 외: 공공기관·비영리·법조·언론 등

AI의 영향을 받는 사람들의 목소리는 빠져 있다

6명

시민사회와 (조금이라도)
관계가 있는 인사

1명

시민 권리를 옹호·감시하는
단체 출신
(김하나 변호사, 민변)

0명

노동조합·인권단체 경력자

+ 성별 균형도 고려되지 않았다(35명 중 7명_20%) (109명 중 18명_16%)_사진보며 센 거라 부정확할 수 있음

“인공지능의 영향을 직접적으로 받는 다양한 사회 집단과 시민사회단체의 참여가 사실상 배제”

2025. 9. 9. 공동성명 「시민사회와 영향받는 사람 배제한 국가인공지능전략위원회, 다시 구성하라」 — 참여연대·민변·진보네트워크센터 등 5개 단체

문제제기 이후 개편했으나 ‘그 나물에 그 밥’

2026. 3 개편: 숫자만 늘었다

민간·분과위원	82 → 127명
대통령 위촉 민간위원	33 → 37명
부위원장 위촉 분과위원	49 → 90명
조직	8분과 → 10분과·2특위·1TF

AI민주주의 분과 신설(위원장 김의영 서울대 교수). 대통령 면담(9월경) 뒤 추천했지만 시민사회 추천 인사 공식 추가 위촉은 아직 없다.

의견수렴: 요식행위로 전락

- 2025.12.16 — 행동계획안 공개
- 2026. 1. 4 — 의견 마감 — 연말연시 약 3주
- 같은 기간 — 시민사회 41쪽 공동의견서 제출
- 2026. 2 — 계획 확정 — 요구 거의 미반영
- 2026. 6경 — AI민주주의 분과 면담 — 1회성

왜 문제인가: 규제 대상이 규제를 심의하는 구조

AI기본법의 목적 (제1조)

국민의 권익과 존엄성 보호,
삶의 질 향상, 국가경쟁력 강화

위원회가 심의·의결하는 것

인공지능 기본계획 수립
인공지능 활용 촉진
고영향 인공지능 규율

그런데 위원회 구성은

규제 대상인 기업 측 인사가 다수
영향받는 사람을 대변할 사람은 없음

이해충돌 소지가 큰 기업 대표·임원이 다수 참여해 직접 정책을 심의 → AI 거버넌스가 잘못 구축되었다는 직접적 증거

민주적 AI 거버넌스를 위한 세 가지 요구

1

투명한 공개와 회피 규정

- 전체 명단·위촉일자·겸직 공개
- 발표 127명과 확인 109명의 차이
- 회의록의 참석자·세부 내용 공개
- 이해충돌 신고·회피 규정 마련

2

당장

시민사회 추천 인사 추가 위촉

- 사회·AI민주주의·데이터 분과에 권리옹호 단체·노조·소비자단체 추천 위원
- 데이터 분과·보안 특위에 개인정보·인권 전문가
- 분과위원은 부위원장 위촉 — 바로 가능

3

시민사회 추천권 제도화

- 운영세칙·시행령에 시민사회 몫 신설
- AI기본법 제7조 4호에 “영향받는 사람의 권익 보호 관련 단체가 추천한 사람” 추가

시민사회는 여전히 위원회 ‘밖’에 있다

100명이 넘는 대통령 직속 법정기구, 민주주의 분과까지 생겼지만 영향받는 사람의 목소리는 아직 안에 없다.

2026년 3월 AI시민행동을 결성한 이유 — 그 목소리를 모아 거버넌스 기구 안으로 전달하는 것.

지금은 조금 더 목소리를 높일 때

