

한국의 AI·데이터 거버넌스는 주권적 관점에서 어떤 과제를 안고 있는가

EG@IG — 'AI 거버넌스 데이터 주권' 연구소

장 소

프란치스코 교육회관 220호

발 제 자

고성혁 · 김나연 · 김민진 · 도은서 · 정다현



전체 세션 구성

발표 · 전문가 패널 토론 · 청중 토론으로 이어지는 다섯 단계 구성

0

도입

AI 주권의 개념과 현황 진단

발표자 1

1

파트 1 · 정당성

민주적 정당성과 시민참여 → 파트 1 패널 반응 (이효진 · 오병일)

발표자 2

2

파트 2 · 실질 내용

데이터 자기결정권과 설명권 → 파트 2 패널 반응 (권은정 · 김법연)

발표자 3

3

전체 패널 / 청중 토론

사회자 질문 · 자유 토론

사회자 · 패널

4

정리 · 제언

한국 AI·데이터 거버넌스의 발전 방향

발표자 1

INTRO

00

국민주권론적 AI 주권의 설정

AI 주권을 국민주권의 관점에서 재정의하고,
두 기준(정당성 · 실질 내용)에 따라 한국의 AI·데이터 거버넌스를 검토한다.

주권의 재정의

·
정당성

·
실질 내용





AI 경쟁은 '누가 질서를 설계하는가'의 문제로 이동한다

AI 경쟁의 핵심은 더 우수한 모델의 개발을 넘어, AI·데이터 질서를 누가 설계·통제하고 그 권력이 누구에게 정당화되는가로 이동하고 있다. 그 결과 '**AI 주권(AI sovereignty)**' 개념이 부상했으나, 아직 학술적·정책적 합의는 존재하지 않는다.



기술적 자립으로 보는 관점

- 자국산 AI 모델 보유
- 컴퓨팅 인프라 · 데이터센터 확보
- 반도체 · 클라우드 등 자원의 자립



규범 · 통제 구조로 보는 관점

- AI를 둘러싼 규범의 형성
- AI · 데이터에 대한 통제 구조
- 그 권력의 정당성 문제

본 세션은 **AI 주권**을 단순한 기술적 자립이 아니라, 국민주권론에 기초한 정치·법철학적 개념으로 재구성한다.



주권 개념의 전개 : 국가에서 국민으로, 다시 디지털·AI로



01

근대 주권

국가가 영토·구성원에 갖는
최종적·배타적 결정권

*Bodin · Hobbes;
베스트팔리아 체제*



02

국민 주권

모든 공권력의 정당성은
국민으로부터 나온다

*Rousseau;
헌법 §1②*



03

디지털 주권

데이터·인프라에 대한 국가 통제와
시민의 데이터 자기결정권



04

AI 주권

AI가 예측·분류·판단으로
의사결정 구조 자체에 개입

주권의 초점은 '국가가 무엇을 통제하는가'에서 '국민이 무엇을 위임했고 어떻게 자기지배를 실현하는가'로 되돌아온다.



근대주권에서 국민주권으로: 정당성 원천의 전환

근대 주권론이 '국가의 최종적 결정 권한'에 주목했다면,
국민주권론은 모든 공권력의 정당성이 궁극적으로 국민으로부터 비롯된다는 점을 강조한다.



근대 주권

- 국가가 외부 간섭 없이 영토·구성원에 대해 갖는 최종적·배타적 결정권.
- 보댕: 국가의 절대적·영속적 권력
홉스: 사회계약으로 형성된 절대 통치권.
- 베스트팔리아 체제 이후 국제질서의 핵심 원리로 자리 잡음.



국민 주권

- 국가가 아니라 국민이 주권의 궁극적 원천이라는 원리로 전환.
- 루소: 인민주권은 일반의지에 기초. 현대 헌정질서의 기본 원리.
- 주권은 국민으로부터 위임된 권력이자, 국민의 자유·권리를 실현하는 헌법적 구조.

헌법 제1조 제2항 “대한민국의 주권은 국민에게 있고, 모든 권력은 국민으로부터 나온다.”



AI의 차별성 : 정보 흐름을 넘어 의사결정 구조로



디지털 주권

- 플랫폼이 정보 환경·소비·정치 담론에 국가 못지않은 영향력 행사.
- 데이터의 저장·이동·통제 등 정보 흐름의 문제를 중심으로 전개.
- 국가의 인프라 통제 + 시민의 데이터 자기결정권, 두 의미로 사용.

※ 개인정보자기결정권에서 확장적 개념으로서 데이터 자기결정권



AI 주권

- AI는 데이터를 기반으로 **예측·분류·판단·추천**을 수행.
- 정보 질서를 구성하는 데 그치지 않고 **의사결정 구조 자체를 매개·재구성**.
- '누가 설계·통제하는가'가 기술 경쟁이 아니라 **권력 배분의 문제로 전환**.



구조적 집중성: AI 특유의 권력 집중

막대한 자원 요구

대규모 AI 개발·운영에는
연산·데이터·반도체·클라우드·연구 역량이 필요

소수에 집중

이 핵심 자원이
소수 국가·글로벌 빅테크에 집중되어 있음

플랫폼 종속

파운데이션 모델·클라우드·API 중심으로
강한 플랫폼 종속 구조 형성



AI는 의사결정에 개입하고, 그 힘은 소수에 집중된다



AI의 특수성

- 데이터 기반 예측·분류·판단·추천으로 의사결정 구조 자체를 매개·재구성한다.
- 대규모 AI는 연산·데이터·반도체·클라우드 등 막대한 자원을 요구한다.
- 자원이 소수 국가·빅테크에 집중되고, 파운데이션 모델·API 중심의 플랫폼 종속이 형성된다.



국민주권론적 재구성

- 국민은 AI의 핵심 입력값인 데이터를 생산하는 주체이다.
- 동시에 AI의 예측·분류·판단을 직접 받는 당사자이다.
- 국민은 AI의 형성·작동·결과 전 과정에 구조적으로 연결된다.

AI는 기술 경쟁을 넘어 권력 배분의 문제가 된다.

→ AI 거버넌스의 정당성은 국민주권의 관점에서 엄격히 검토되어야 한다.



국민주권 관점의 AI 주권 : 두 가지 검토 기준

기준 1 · 정당성



AI 거버넌스는 국민으로부터 정당하게 위임받은 권한 구조를 갖추고 있는가?

민주적 정당성과 시민참여의 실질성 → 파트 1

기준 2 · 실질 내용



AI 거버넌스는 위임의 실질 내용인 국민의 기본권·자기결정권을 보장하고 있는가?

데이터 자기결정권과 설명권 → 파트 2

핵심 질문

현재 우리의 AI 거버넌스 체계는 국민주권론적 관점에서 정당한가?

PART

01

정당성

민주적 정당성과 시민참여의 실질성으로 위임 구조를 검토한다.

민주적 정당성

· 시민참여





정당성 = 위임 구조의 적법성 × 시민 참여의 실질성

국민주권의 실질적 측면에서 정당성은 '권한이 국민으로부터 적법하게 위임·행사되는가'만으로 완성되지 않는다.
주권자가 자가지배의 주체가 되려면 '영향받는 시민이 직접 참여하는가'까지 물어야 한다.



쟁점 1 · 민주적 정당성

형식적 측면

권한이 국민 → 국회 → 법률 → 행정기구로 적법하게 위임되었는가

검토 기준 : 법률 기반 · 권한 분산 · 책임성 · 투명성



쟁점 2 · 시민참여

실질적 측면

영향받는 시민이 정책 원칙 형성 과정에 직접 참여하는가

검토 기준 : 참여 주체 · 선발 방식 · 정책 반영 효력



제도적 기반 이후의 과제: 권한 행사의 방식

핵심 주장

국민주권적 AI 주권의 관점에서, 한국은 민주적 정당성의 제도적 기반(input)을 일정 부분 확보했다. 그러나 정당성의 실질화는 권한 행사의 방식, 즉 책임성·투명성·참여성(throughput)에 달려 있다. 핵심은 직접참여의 확대 그 자체가 아니라, 위임·자문형 구조를 어떻게 작동시키고 보완할 것인가이다.



제도적 기반 (input)

- 법률 기반 위임 : 헌법 §1② → 국회·입법 → AI기본법(2026.1.22)
→ 행정 위임(전략위·주관 부처)
- 포괄적·수평적 법체계 + 법정 조정 장치로 정당성의 제도적 출발점 확보



작동 방식 (throughput)

관건

- 권한이 실제로 어떻게 행사되는가 : 책임성 · 투명성 · 참여성
- 법적 근거의 존재가 곧 정당성의 완성은 아니다.
절차적 정당성으로 검증 필요 (Bovens 2006)

이어서 근거 ① 법률 기반 위임(input)과 근거 ② 절차적 정당성(throughput)을 차례로 검토한다.



한국 AI 거버넌스 체계 : 위임·집행의 계층 구조

국민 (주권)

헌법 §1②

→ '주권은 국민에게, 모든 권력은 국민으로부터'. 모든 권한의 원천.

국회 · 입법 토대

AI기본법(2026.1.22 시행)

→ AI 전략·진흥·거버넌스 + 윤리원칙·투명성·안전성·고영향 AI 관리 등 신뢰성 확보

정점 조정 · 주관

국가AI전략위원회(대통령 직속·심의/조정) · 과학기술정보통신부(부총리 격상·주관/집행)

분과 · 전문기관

분과위(AI민주주의 분과 등)·협의체 · AI안전연구소·정책센터 · (독립) 개인정보보호위원회

권한 분산 전략위(조정) – 과기정통부(주관·집행) – 관계부처/분과(분야별) – 전문기관(지원·집행)

시민 진입점



- 위촉 민간위원
- 입법예고 의견수렴

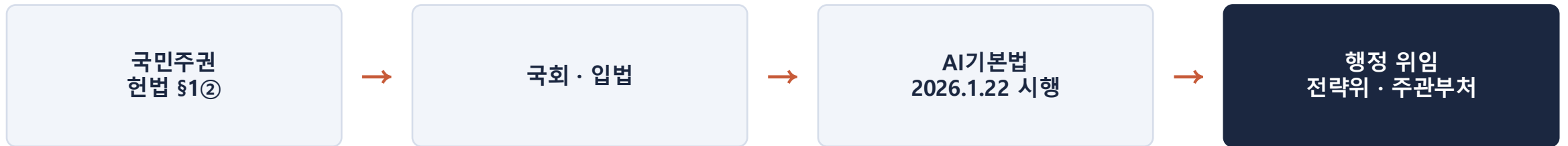
→ 간접(위임형) 참여

무작위 속의 등 직접 참여 기제는 별도로 제도화되어 있지 않다.



법률 기반 위임 : 정당성의 제도적 출발점

민주적 정당성은 국가가 권한을 행사한다는 사실만으로 확보되지 않고,
그 권한이 '어떤 규칙에 근거해 누구로부터 위임되었는가'로 설명된다. (Beetham 1991 · Schmidt 2013)



포괄적·수평적 법체계

- **AI기본법**은 개별 분야법이 아니라 AI 전반의 기본 원칙·추진체계를 정하는 수평적·포괄적 법체계다.
- **제4장**에서 윤리원칙·투명성·안전성·고영향 AI 관리 등 신뢰성 확보 장치를 별도로 둔다.



법정 조정 장치

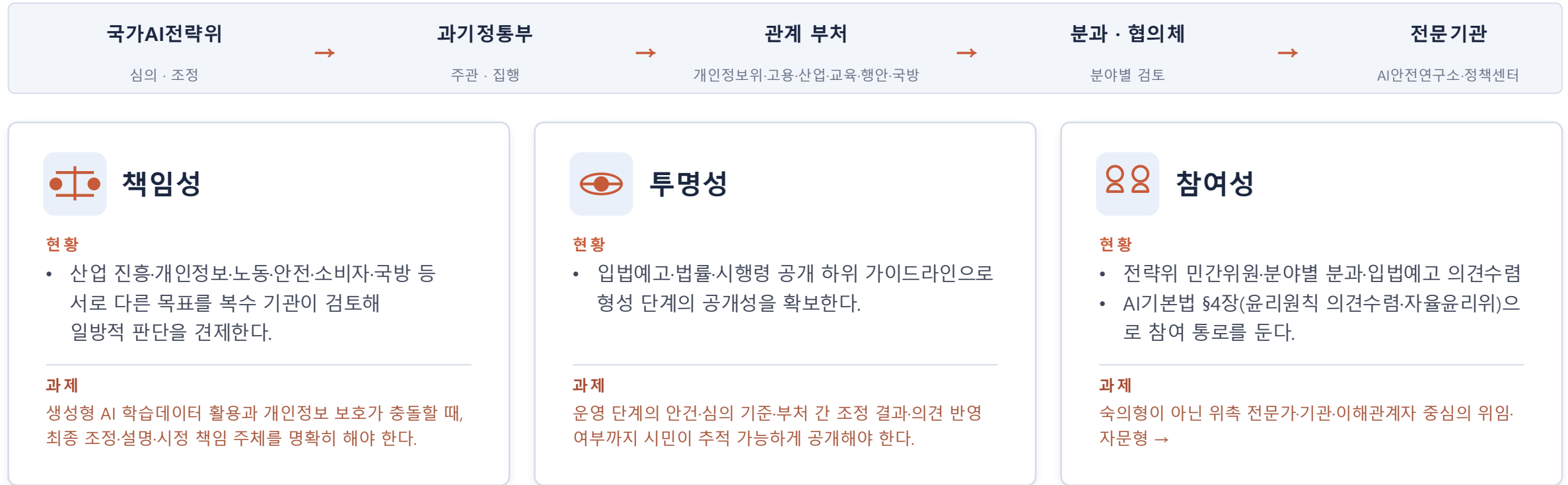
- **국가AI전략위원회**는 여러 영역에 분산된 정책을 조정·종합하는 법정 장치다.



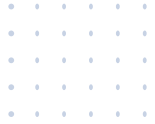
권한 분산의 절차적 정당성 : 책임성 · 투명성 · 참여성

정당성은 위임의 근거(input)뿐 아니라 권한이 어떻게 행사되는가로도 평가된다.

Schmidt & Wood(2019)의 책임성·투명성·포용성·개방성 중 포용성·개방성을 참여성으로 통합해 세 기준으로 본다.



종합 권한 분산은 민주적 정당성의 외형을 보강한다. 실질화하려면 조정 책임 명확화 · 운영 단계 정보공개 · 의견 반영 피드백 · 대표성·이해당사자 균형 기준의 구체화가 필요하다. (집행수단 유예 시 그 기준·절차의 설명 포함)



간접적·위임형 구조를 중심으로 설계된 시민 참여

국가	거버넌스 체계 (정점·법형식)	민주적 정당성	시민 참여
한국	- 포괄법(AI기본법) - 대통령 직속 법정 전략위 - 과기정통부(부총리)	대의·법률 위임 + 민관 참여형 → 정당성의 제도적 기반	간접(위임형)
미국	- 포괄법 부재; 행정명령(EO) - 주(州)법, 정점 기구 부재	행정·시장 기반; 연방-주 충돌로 정당성 다툼	간접(연방 차원 얇음)
중국	- 분야별 규정 - 개정 사이버보안법; 당-국가·CAC	국민주권 아닌 국가(당)주권	형식적
일본	- AI추진법(벌칙 없는 기본법) - 총리 주재 전략본부	대의 기반이나 연성·내각 주도	간접(시민=리터러시)

네 사례 모두 기술·인프라 자립을 추구하는 한편, 시민참여는 대체로 위임형 구조를 중심으로 설계되어 있다. 이에 따라 다음 쟁점에서는 시민참여의 주체와 효력을 비교한다.



주권행사와 시민참여 양식의 설정: 주체/효력, 위임형/직접형

주권 행사의 '주체(위임↔직접)'와 '효력(자문↔형성)' 두 축으로 시민참여를 본다. *Fung 민주주의 큐브 · Arnstein 참여 사다리 · OECD 숙의의 물결*

	주체 →		
효력 ↑	형성 · 구속	위임 (대표·전문가)	형성 · 구속
	자문 · 청취	위임 (대표·전문가)	직접 (당사자 시민)
			자문 · 청취
			직접 (당사자 시민)



위임형 참여

주권의 위임적 행사

- 시민이 권한을 선출직 대표·전문가에게 위임하고, 대의제와 자문·의견수렴(공청회·입법예고)으로 간접 반영하는 방식.
- 시민은 최종 심의·결정에 직접 들어가지 않는다.



직접형 참여

주권의 직접적 행사

- 영향받는 시민 자신이 매개 없이 심의에 들어가는 방식.
- 통상 무작위 추출(sortition) 시민이 정보·숙의를 거쳐 견해·권고를 형성한다(숙의 미니퍼블릭).

유의 '직접형'은 직접민주주의(구속력 있는 국민투표)와 다르다. 직접 참여라도 효력은 '자문 ~ 형성·구속' 스펙트럼이며, 그래서 주체(위임↔직접) × 효력(자문↔형성) 두 축으로 본다.



한국 · 대만 · 벨기에 : 시민참여의 구성 · 효력 비교

한국

위임 · 자문형

- 위촉 민간위원·분과 + 입법예고 의견수렴
- AI기본법 §3⑤: 취약계층 참여·반영 '노력' 의무 (2026.7 시행)
- 무작위 시민 숙의는 미제도화

선발 · 규모

행정부 위촉 중심 · 무작위 선발·대표성 기준 미확인

주체 '위임' / 효력 '자문'

대만

직접 · 형성형 (비구속)

- 2024.3 moda-Stanford DDL '정보무결성' 시민 숙의
- 총화표집 약 447명이 직접 참여
- 플랫폼 자율규제·딥페이크 대응 등 규제 원칙 형성 논의

선발 · 규모

총화 무작위 표집 · 단발성·비구속

주체 '직접' / 효력 '형성(비구속)'

벨기에

직접 · 자문형

- 2024 EU 의장국 AI 시민패널
- 16,200명 초청 → 1,070명 응답 → 60명 무작위 선발
- 3주말 대면 숙의로 9개 '시민 비전' 메시지 작성

선발 · 규모

EU·자국 정책 참고·영감 · 구속력 없음

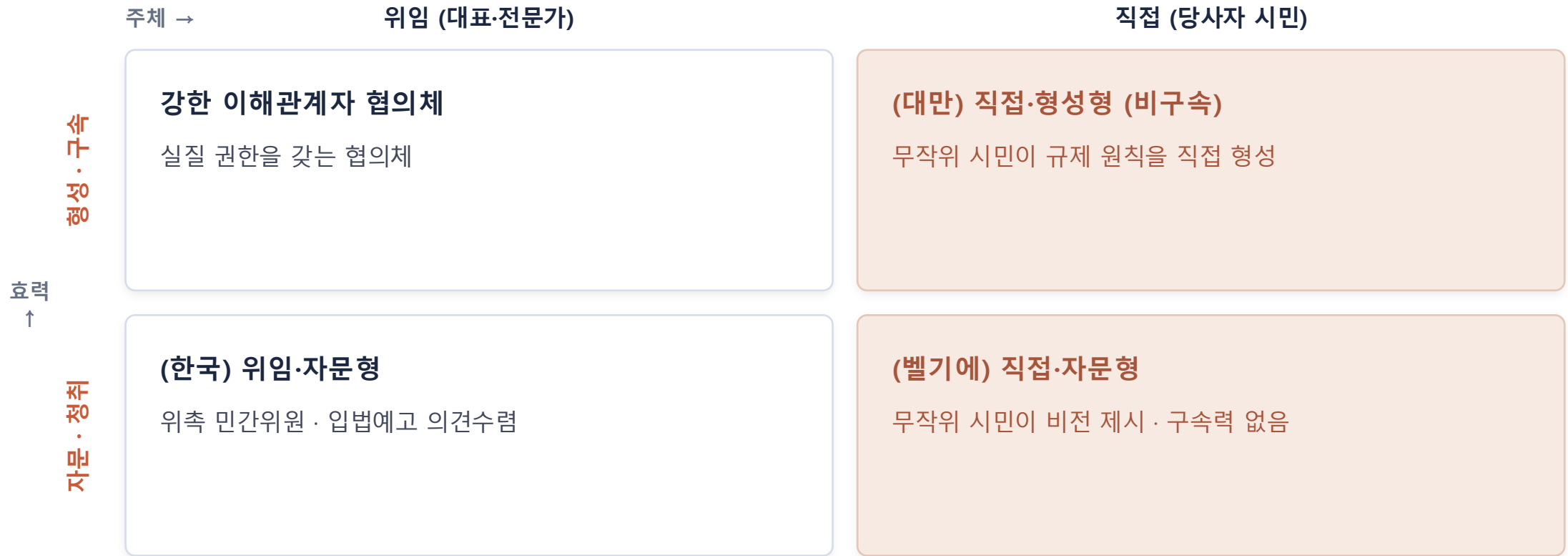
주체 '직접' / 효력 '자문'

두 축 같은 '직접형'도 대만은 규제 원칙을 형성하고(형성형), 벨기에는 자문에 그친다(자문형). 주체(위임↔직접)와 효력(자문↔형성)은 별개의 축이다.



설정된 기준 양식에 국가별 사례 적용

주권 행사의 '주체(위임↔직접)'와 '효력(자문↔형성)' 두 축으로 시민참여를 본다. *Fung 민주주의 큐브 · Arnstein 참여 사다리 · OECD 속의의 물결*





정당성의 출발점과 남은 쟁점

한국은 대의제, AI기본법, 법정 전략위를 통해 민주적 정당성의 제도적 출발점을 마련했다.

그러나 정당성이 실질화되기 위해서는 세 가지 질문이 남는다. 위임만으로 충분한가, 직접 참여는 필요한가, 그 참여의 효력은 어디까지 인정할 것인가.



01

진단 (공통점)

한·미·중·일모두 기술·인프라 자립은 강조하지만, 시민참여는 주로 위임형 구조를 중심으로 설계되어 있다.. 핵심은 직접참여의 확대 그 자체가 아니라, 위임된 권한이 어떤 절차와 책임 구조 속에서 행사되는가이다.



02

한국의 위치

법률 기반 위임으로 절차적 정당성의 제도적 기반은 확보. 다만 throughput(책임성·투명성·참여성) 중 '참여성'은 위임형에 집중되어 있다.



03

남은 쟁점

대만과 벨기에는 직접참여가 위임 구조를 대체하기보다 보완할 수 있음을 보여준다. 쟁점은 참여를 언제 필요로 볼 것인지, 그리고 그 결과를 자문적 의견으로 돌지 정책 형성 과정에 반영할지에 있다.

파트 2 정당성의 '출발점'은 마련됐다. 위임형 참여만으로 실질적 정당성이 충분한지 데이터 자기결정권과 설명권의 '실질 내용'에서 이어 검토한다.



패널 지정 질문

한국은 법률 기반 위임으로 민주적 정당성의 출발점을 마련했다.

다만 그 정당성이 실질화되려면 위임만으로 충분한지, 직접 참여가 필요한지, 그 효력을 어디까지 설계할지가 남는다.

Q1

위임의 충분성

- 대의제와 전문가 위원회를 통한 위임형 참여만으로 민주적 정당성이 충분한가?
- 한국이 모범처럼 보이는 이유는 제도적 외형 때문인가, 실제 반영 때문인가?

Q2

직접 참여의 필요성

- AI의 전문성·변화 속도를 고려할 때 무작위 시민 속의·시민의회는 필요한 보완인가?
- 아니면 형식적 참여가 전문적 판단·책임을 흐릴 위험이 더 큰가?

Q3

직접 참여의 효력

- 직접 참여가 필요하다면 그 결과는 법정 전략위 구조에서 어디까지 효력을 가져야 하는가?
- 단순 자문, 원칙 형성 참고, 공식 안건 상정·답변 의무 중 어디인가?

PART

02

실질 내용

위임의 실질인 기본권을 보장하는가? 데이터 자기결정권과 설명권을 사례로 검토한다.

데이터 자기결정권

설명권





실질 내용의 두 축: 데이터 자기결정권과 설명권

그 권력이 국민의 자유·권리를 실질적으로 보호하는 방식으로 작동하는지가 국민주권적 AI 주권의 핵심 쟁점이 된다.
실질 내용은 국민이 자신의 데이터와 AI 결정에 대해 실제로 통제권·설명요구권을 행사할 수 있는가로 검토한다.



쟁점 1 · 데이터 자기결정권

정보주체는 자신의 데이터가 언제·어디서·어떻게 처리·활용되는지
인지하고 통제할 수 있는가

무엇을 보는가

- 수집·이용·이전·삭제 전 단계의 통제 (개인정보보호법 §15·18·35~37·28의8)
- 비식별·공개 데이터의 보호 공백
- 학습 후 가중치 내재화(암기)로 인한 삭제권의 기술적 한계



쟁점 2 · 설명권

AI가 내린 판단의 이유를 이해하고 그 결과를 다룰 수 있는가

무엇을 보는가

- 고영향 AI의 설명가능성 (AI기본법 §31·34·35, 개인정보보호법 §37의2)
- 고영향 기준의 모호성과 규제 우회 여지
- 블랙박스·해외 API 종속에 따른 설명 책임 공백

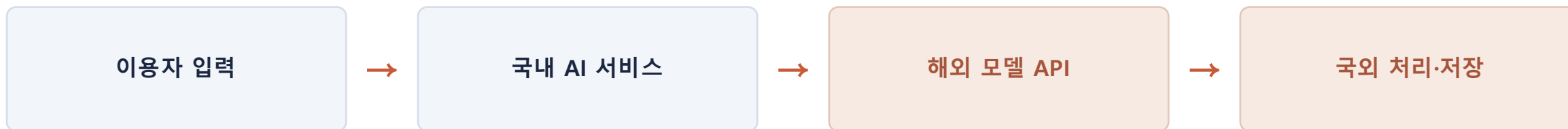
검토 방식 법령 조문의 나열이 아니라, Meta·OpenAI 사례와 해외 API 종속 현상을 통해 규범적 강점과 실무적 공백을 함께 본다.



AI 학습 과정에서 정보주체의 데이터 통제권 보장

핵심 주장

다수 국내 AI 서비스가 OpenAI·Anthropic·Google의 API에 의존한다. 이 경우 이용자 데이터가 국경을 넘어 해외 모델 제공자 시스템에서 처리·저장될 수 있어, 이용자가 자신의 데이터가 어디서 누구에 의해 처리되는지 실질적으로 통제하기 어렵다.



국민주권적 차원의 문제

데이터 처리 경로가 다층적으로 분산될수록, 국가는 AI 기술을 활용할 수 있더라도 국민의 데이터 통제권이 충분히 보장되기는 어렵다. 이러한 점에서 국민주권적 AI 주권 실질화의 과제를 남기고 있다.



OpenAI와 Meta 데이터 수집 논란

OpenAI

2023 · 개인정보보호위원회

- 개인정보 유출 인지 후 24시간 내 신고 의무 위반
→ 과태료 360만 원 처분
- 안전조치의무 위반은 인정되지 않았으나, 처리방침 영문 제공·동의 절차 미비 등 개인정보 처리의 투명성 문제가 지적됨
- **핵심 쟁점**
AI 서비스 이용자는 자신의 개인정보 처리 과정을 충분히 인지하고, 실질적으로 통제할 수 있는가?

Meta

2024 · EU 규제기관·시민사회

- Facebook·Instagram의 공개 콘텐츠를 AI 학습에 활용하려는 정책 추진
- 이용자의 실질적 동의·통제권 보장이 부족하다는 비판
→ 유럽 규제기관·시민사회 반발
- **핵심 쟁점**
공개 콘텐츠는 AI 학습에 자유롭게 활용될 수 있는가?

AI 서비스 이용자는 자신의 데이터 처리·활용 과정을 충분히 알고, 이를 실질적으로 통제할 수 있는가?



현행 체계는 기본적 보호 장치를 갖추고 있다

근거 조항	보장 내용
개인정보 보호법 §15 · §18	수집·이용의 적법 근거 요구, 목적 외 이용·제3자 제공 제한 → 학습 활용 시 별도 근거 필요
§35 ~ §37	정보주체의 열람권 · 정정·삭제권 · 처리정지권 → 수집 이후에도 사후 통제 가능
§28외8	개인정보 국외 이전에 별도 요건 부과



개인정보위 「생성형 AI 개인정보 처리 안내서」 (2025)

Enterprise Terms-DPA 계약으로 입력 데이터 소유권·재이용 제한·파기 조건 명확화,
API 접근통제·입출력 필터링·시스템 보안·지속 벤치마크 권고함.

→ 데이터가 무분별하게 수집·이용·이전되는 것을 방지하는 기본 보호 장치는 갖췄다고 평가할 수 있다.



법적 권리와 실질적 행사 사이의 간극

① 보호 범위의 공백

- 현행법은 '개인정보'(식별 가능성) 중심으로 설계된다(§2).
- AI 학습에는 비식별·크롤링 데이터가 광범위하게 쓰인다.
- **현행 개인정보 보호법 체계에서는 개인정보가 아닌 데이터에 대해 자기결정권을 직접 행사하기 어렵다** (공개 게시물·블로그·이미지 등).
- 이 영역은 저작권법·계약·구속력 없는 가이드라인에 의존 → **규율 공백**.

② 기술적 공백

- 법은 열람·정정·삭제·처리정지권을 보장한다(§35~37).
- 그러나 **생성형 AI에서 데이터는 학습 후 모델 weight에 내재화**된다.
- '암기(memorization)'·영구 기억으로 원본이 그대로 남지 않을 수 있다.
→ 법은 삭제권을 보장하나 AI 환경에서는 삭제권의 실질적 이행과 검증이 불명확해질 수 있다.

핵심은 권리가 선언되어 있는가가 아니라, 국민이 실제로 자신의 데이터에 대해 실질적 통제 권을 행사할 수 있는가이다.



AI 판단의 설명가능성과 이의제기권

핵심 주장

삶에 직결되는 '고영향 AI'에서는 사용 사실의 고지를 넘어, **판단의 이유를 이해하고 그 결과를 다룰 수 있는 설명가능성**이 주권의 핵심이다.



왜 설명권인가 · 고영향 AI

채용

서류·면접 자동 심사의 탈락 사유

금융

대출·한도 거절 판단의 근거

복지

수급 자격·중단 결정의 이유

→ '사용 사실 고지'만으로는 부족하다. 판단의 이유를 이해·반박할 수 있어야 실질적 권리가 된다.



사례 · 해외 API 종속으로 인한 설명 불가

- 국내 공공기관·기업이 자체 모델 대신 OpenAI·Google 등 해외 LLM API로 고영향 서비스를 구축한다.
- 차별적 결과에 이의를 제기해도 국내 배포자가 해외 API의 내부 구조를 충분히 설명하기 어렵다는 점에서, 설명 책임이 국내 배포자와 원천 모델 공급자 사이에서 불명확해질 수 있다.
- 글로벌 원천 모델 공급자의 설명 의무가 한국 이용자에게 어디까지 미치는지는 여전히 불명확하다
→ 입증 책임과 피해는 시민의 몫으로 남는다.

핵심 설명권이 충분히 보장되지 않는 고영향 AI에서는 '고지'가 있더라도 실질적 이의제기가 어려운 절차적 공백이 남을 수 있다.



설명 의무의 법제화와 고영향 AI 규율



설명 · 투명성 의무

- **AI기본법 §31·§34·§35**
: 사용 사실 고지, 설명 방안 수립·게시, 안전성·신뢰성 확보 의무
- **개인정보 보호법 §37의2**
: 완전 자동화 결정에 대한 정보주체의 설명요구권 근거



고영향 AI의 차등 규제

- **AI기본법 §2④**
: 생명·신체·기본권에 중대한 영향을 주는 영역을 '고영향 AI'로 정의
- 의료·에너지·교통·채용·생체인식 등 11개 핵심 분야에
강한 조건을 차등 부과

강점 설명가능성을 법적 의제로 전면화하여 결정 근거 제공·데이터 투명성·이용자 보호를 의무화하고, 기본권 중대 영역을 차등 규제하려 했다는 점에서 분명한 진전이다.



설명 책임에 대한 공백이 발생할 수 있는 세 지점

현행법은 고영향 AI에 설명 의무를 부과했지만, 기준 · 기술 · 공급망 세 지점에서 책임이 흩어져 시민의 설명요구권이 형식화될 수 있다.

1

기준

고영향 기준의 모호성 · 규제 우회

고영향 해당 여부를 사업자가 1차적으로 스스로 판단 → 이용자 기본권보다 사업자 편익이 우선될 여지.
설명 의무가 일반적 공시에 그쳐 개별 이의제기·인간 재검토로 이어지지 못한다.

2

기술

설명가능성의 기술적 복잡성

수백만 가중치의 신경망은 구조상 블랙박스적 성격이 강하다.
배포자조차 내부 로직을 몰라 제공자와 이용자 모두가 납득 가능한 설명을 확보하기 어려운 공백이 생길 수 있다.

3

공급망

해외 API 종속과 책임 회피

국내 배포자가 해외 API를 커스텀 → 공급망이 다층화되어 데이터 탃·모델 탃이 불분명.

귀결 설명 책임이 명확히 배분되지 않으면, 시민의 설명요구권은 실질적으로 행사되기 어려워질 가능성이 존재한다.



선언된 권리, 미완의 통제와 설명

현행법은 데이터 자기결정권과 설명권을 명문화하고 고영향 AI 투명성 책무를 선언했다.
그러나 **보호 범위·기술적 공백·해외 API 종속**으로 인해, 권리의 '선언'과 실질적 '행사' 사이의 간극이 핵심 과제로 남는다.



01

쟁점 1 · 데이터 자기결정권

권리는 명문화됐으나

- ① 식별성 중심 '개인정보'로 보호 범위가 한정되고,
- ② 데이터가 가중치로 내재화·암기되어 학습 이후에는 원본 삭제와 그 효과를 확인하기 어렵다.



02

쟁점 2 · 설명권

고영향 AI 투명성 책무가 법적 의제로 전면화됐으나,
설명 책임 소재가 불명확해지는 구조 및 기술적으로 설명하기 어렵다는 어려움이 남는다.



03

공통 합의 · 해외 API 종속

국내 서비스가 글로벌 원천 모델에 종속되며 책임·설명 공백이 발생. **원천 공급자와 국내 배포자 간 책임 규율**이 실질화의 관건이다.

종합 파트 1의 '정당성(누가 정하는가)'과 파트 2의 '실질 내용(무엇이 담기는가)'이 함께 충족될 때,
AI·데이터 질서는 비로소 '국민주권적 AI 주권'에 다가선다.



지정 패널 질문

현행법은 권리를 명문화하고 고영향 AI 투명성 책무를 선언했다.

그러나 개인정보 중심의 보호 범위, weight 내재화의 기술적 공백, 해외 API 종속으로 인한 책임 배분의 불명확성 때문에 실질적 통제·설명은 여전히 과제로 남아 있다.

Q1

보호 범위 확장

식별성 중심 '개인정보'를 넘어, AI 학습에 대량 활용되는 비식별·공개 데이터까지 정보보호 패러다임을 전체 데이터 생태계로 확장하기 위한 과제는 무엇인가?

Q2

기술적 삭제권 표준

데이터가 가중치로 내재화·암기되어 원본 삭제와 그 효과 확인이 어려운 한계에서, 사후 통제권을 보장할 기술적 표준·보안 통제 장치는 무엇이어야 하는가?

Q3

해외 API 책임 규율

국내 서비스가 글로벌 API에 종속되며 설명 책임 공백이 생긴다.
해외 원천 모델 공급자와 국내 배포자 간 책임을 실효성 있게 규율할 정책 접근이 존재하는가?

DISCUSSION

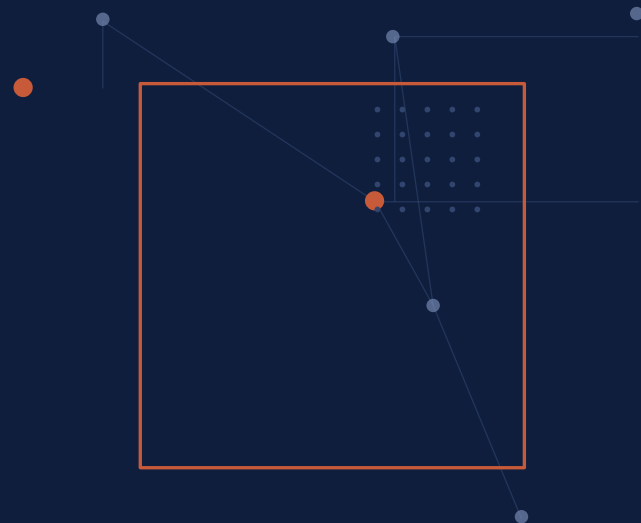
03

전체 플로어 질문

오늘의 논의를 모아 패널 전체 청중과 함께 좁혀가는 두 개의 핵심 질문.

실무·법제 관점

시민 참여 관점





전체 플로어 질문

오늘의 논의를 모아 패널 전체와 청중에게 던지는 두 개의 핵심 질문

1 실무 · 법제 관점

국민주권적 AI 주권을 실질적으로 구현하려면,
한국 AI 거버넌스는 무엇을 가장 우선 보완해야 하는가?
시민참여 확대인가, 권리 보장의 실효성 강화인가,
새로운 제도적 거버넌스 설계인가?

논의 배경

파트 1은 정당성(위임이 충분한가), 파트 2는 실질 내용(권리가 실효적인가)을
짚었다. 두 축 중 어디를 먼저 보완하느냐가 제도 설계의 우선순위를 가른다.

2 시민 참여 관점

'국민주권론적 AI 주권'이라는 새로운 개념의 도입이
현 시대에 필요한 논의인가? 시민을 AI 질서의 공동 형성자로 세우
는 참여 설계는 어디까지 가능한가?

논의 배경

기존 논의는 AI 주권을 '기술 자립'으로 보아 시민을 소비자·수혜자로 두었다.
국민주권적 재정의는 시민을 AI 질서 형성 과정의 참여 주체로 위치 지우는
전환이다. 이 전환이 지금 필요한지, 어떻게 설계할지가 관건이다.

CLOSING

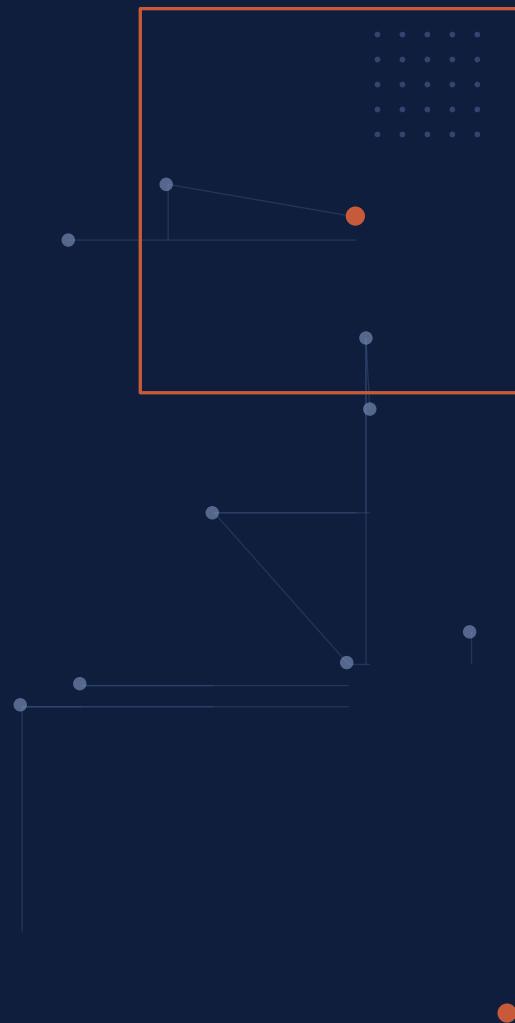
04

정리 및 제언

전체 요약

해외 사례 종합

최종 제언





지금까지의 논의

핵심 질문 현재 우리의 AI 거버넌스 체계는 국민주권론적 관점에서 정당한가?

도입



국민주권적 AI 주권

AI 주권을 기술 자립이 아니라 국민으로부터 위임된 권력이자 자유·권리 실현 구조로 재정의. (1) 정당성 (2) 실질 내용의 이중 기준.

파트 1



정당성

포괄법·대통령 직속 전략위·부총리 주관부처로 정당성의 제도적 기반은 확보. 다만 시민참여는 위임·자문형 구조를 중심으로 이루어진다.

파트 2



실질 내용

개인정보보호법·AI기본법이 데이터 자기결정권·설명권을 명문화. 다만 보호범위·기술적 삭제·해외 API 종속의 공백이 남는다.

권리를 선언하는 속도와 그것을 실질화하는 제도를 갖추는 속도 사이의 간극.



해외 사례의 시사점: 참여·통제·책임

축 (출처)	핵심 해외 사례	한국에의 제언
민주적 정당성·시민참여 (파트 1)	- 대만 정보무결성 숙의(2024, 약 447명) - 벨기에 AI 시민패널(2024, 60명)	위임·자문형 구조를 보완할 직접·숙의형 참여 통로 검토 (효력 설계는 열린 과제)
데이터 자기결정권 (파트 2)	- OpenAI 과태료·개선 권고(2023) - Meta 공개 콘텐츠 학습 논란(2024)	- 비식별·공개 데이터까지 보호 확장 - 기술적 삭제권 표준 마련
설명권 (파트 2)	해외 LLM API 종속 → 설명 불가·책임 회피 구조	- 해외 원천모델-국내 배포자 책임 분담 - 역외 실효성 제도화

종합 해외 사례는 수입 모델이 아니라, 한국 맥락에서 '직접 참여 · 데이터 통제 · 설명 책임'을 함께 설계하라는 방향을 가리킨다.

국민주권적 AI 주권의 실질화를 향하여

AI 주권은 규제 법률이나 조정 기구의 존재만으로 완성되지 않는다.
국민이 실제로 자신의 데이터와 AI의 결정에 대해 통제권·설명요구권을 행사할 수 있을 때 실질화된다.



정당성의 실질화

위임 구조를 전제로 하되, 직접·숙의형 시민참여를 보완하고
운영 단계 정보공개와 의견 반영 피드백을 제도화한다.



기본권의 실질 보장

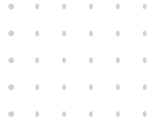
개인정보를 넘어선 데이터 생태계 규율, 기술적 삭제·추론경로 추적, 해외 API 설명책임 분담을 갖춘다.



시민을 주권자로

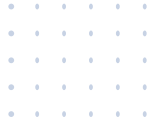
시민을 정책의 소비자·수혜자에 그치지 않고, AI 질서 형성에 참여할 수 있는 주체로 위치 지운다.

비전 선언된 권리가 일상에서 실제로 행사되는, 국민이 실질적 주체로 참여하는 국민주권적 AI 주권.



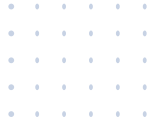
참고문헌

- [1] 소버린 AI가 뭐길래? 데이터 주권을 넘어서 AI 주권의 모든 것, 한국딥러닝, 2025, <https://www.koreadeep.com/blog/sovereign-ai>.
- [2] 이현경, 주권 개념의 본질과 근대주권론에 대한 비판적 고찰 - 법철학적 주권론을 향하여 -, 이화여자대학교 법학논집 제27권 제2호 통권 80호, 2022.
- [3] 김성주, 주권개념의 역사적 변천과 국제사회로의 투영, 한국정치외교사논총 vol.27, no.2, pp. 195-224(30), 한국정치외교사학회, 2006.
- [4] 이계일, 주권론의 위협상황과 국민주권론의 재구성가능성에 대한 고찰 - 이 시대 주권론은 가능한가? 가능하다면 어떠한 조건 하에서? -, 법철학연구 제16권 제1호: 233~290, 한국법철학회, 2013.
- [5] 대한민국 헌법 제1조 제2항
- [6] 이희옥, 인공지능 주권 논의와 데이터 법·정책에 관한 비판적 고찰, 2024.
- [7] Leonardo Gambacorta & Vatsala Shreeti, Big Techs' AI Empire, VOXEU, 2025, <https://cepr.org/voxeu/columns/big-techs-ai-empire>.
- [8] 정한영, 스탠퍼드 HAI, 'AI 인데스 2026'...'생성형 AI, 이제 기술이 아니라 지배 인프라', 산업·과학·국가 시스템까지 재편, 인공지능신문 2026, <https://www.aitimes.kr/news/articleView.html?idxno=39542>.
- [9] David Beetham, The Legitimation of Power (Macmillan, 1991).
- [10] V. A. Schmidt, "Democracy and Legitimacy in the EU Revisited: Input, Output and 'Throughput'," Political Studies 61(1), 2013.
- [11] 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」 제1조 및 제4장(인공지능윤리 및 신뢰성 확보: 윤리원칙, 민간자율인공지능윤리위원회, 투명성·안전성 확보 의무, 고영향 인공지능 관련 책무 등), 국가법령정보센터.
- [12] AI기본법 제7·8조(국가AI전략위) 및 「국가인공지능전략위원회 설치·운영 규정」; 전략위 법정기구 출범(2026.1.22.).
- [13] Mark Bovens, "Analysing and Assessing Public Accountability," European Governance Papers C-06-01, 2006.
- [14] V. A. Schmidt & M. Wood, "Conceptualizing Throughput Legitimacy," Public Administration 97(4), 2019 — 책임성·투명성·포용성·개방성.
- [15] 개인정보보호위원회, 「생성형 AI 개발·활용을 위한 개인정보 처리 안내서」(2025).
- [16] 「국가인공지능전략위원회 설치·운영 규정」 제정령안 입법예고.
- [17] A. Fung, "Varieties of Participation in Complex Governance," Public Administration Review 66(s1), 2006 — 참여자 선발·권한('민주주의 큐브').
- [18] A. Meijer, D. Curtin & M. Hillebrandt, "Open Government: Connecting Vision and Voice," Int'l Review of Administrative Sciences 78(1), 2012.
- [19] 국가인공지능전략위원회 출범(2025.9.8., 민간위원 34명·8개 분과) 및 2026.2~3월 확대 개편(총 127명, 과학·인재 분과 분리, AI민주주의 분과 신설, 보안·지역 특위·법률TF); 민간위원 위촉 기준 부재(인공지능산업법학회 의견서, 2025.12). aikorea.go.kr; 데일리시큐; 법률신문.
- [20] 고영향 AI 과태료·사실조사 유예(계도기간 최소 1년 이상, 2026.6 현재 운영). 경향신문(2025.9.); 법률사무소 변화(2026.6.).
- [21] 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법」, 제3조 제5항. 동 조항은 2026.1.20. 신설되어 2026.7.21. 시행 예정이며, 국가와 지방자치단체가 AI 정책의 개발·수립 과정에서 장애인·고령자 등 AI취약계층의 참여를 보장하고 의견이 반영될 수 있도록 노력할 의무를 규정한다.
- [22] 미국 — 포괄 연방 AI법 부재; 행정명령 EO 14179·14365·2026.6, 주법 우선적용 시도. White House; Paul Hastings; White & Case.



참고문헌

- [23] 일본 — AI추진법(2025.5. 제정, 9.1. 시행, 벌칙 없는 기본법), 총리 주재 AI전략본부. White & Case; FPF; CSIS.
- [24] 중국 — 포괄 AI법 부재이나 2026년 입법의제로 추진 중, 분야별 규정 + 개정 사이버보안법(2026.1. 시행), 당-국가 거버넌스. East Asia Forum; The Diplomat.
- [25] S. R. Arnstein, "A Ladder of Citizen Participation," Journal of the American Institute of Planners 35(4), 1969.
- [26] (보조) IAP2, Spectrum of Public Participation — Inform-Consult-Involve-Collaborate-Empower.
- [27] OECD, Catching the Deliberative Wave, 2020 — 무작위 선발 숙의기구의 실재·확산.
- [28] 'AI민주주의 분과' 신설(2026.3., 분과위원장 김의영) 및 구성 비판(AI 전문가 부재·여성 7/36명); AI시민행동 간담회(2026.6.25.). ZDNet Korea; 이로운넷; 디지털정의네트워크.
- [29] 문화연대 외 시민사회 공동성명(2025.9.) — '산업계 일색 전략위' 비판, '국민을 AI의 주체가 아니라 소비자·수혜자로 본다'.
- [30] 대만 — moda-CIP Alignment Assemblies(2023~); moda, "Utilizing AI to Promote Information Integrity" 온라인 시민 숙의(2024.3.23., 국립양밍교통대/Stanford Deliberative Democracy Lab 참여, 총화 무작위 표집 약 447명). 숙의 결과는 AI 발전 정책의 참고자료로 제시되었으므로, 본문에서는 '직접·형성형(비구속)'으로 분류한다. cf. 산업육성형 Ideathon은 권리기반 시민숙의 사례에서 제외.
- [31] 벨기에 — Belgian Presidency of the Council of the EU, AI 시민패널(2024): 전국 16,200명 초청, 1,070명 응답, 60명 무작위 선발; 9개 핵심 메시지를 담은 『A Citizens' View of Artificial Intelligence within the EU』 보고서(2024.5.). 보고서는 EU 및 벨기에 정책에 대한 참고·영감 자료로 제시되었으며 법적 구속력은 없다. cf. 2019년 출범한 Ostbelgien 상설 시민대화는 벨기에의 일반 숙의민주주의 제도 기반이나 AI 전용 거버넌스는 아니다.
- [32] (이론) J.-J. Rousseau, 『사회계약론』(주권의 양도·대표 불가); H. Landemore, Open Democracy(2020); 송경호, 「AI 이후의 민주주의: 참여의 역설」(2025).
- [33] (참고) One Project, "How to Make AI Serve the Public"(2026); People Powered, "Impacts of Citizens' Assemblies"(2026).
- [34] 개인정보보호위원회, '챗GPT 개발사 OpenAI에 과태료 부과 및 개선 권고', 2023 .
- [35] 개인정보보호위원회, 2024년 모니터링 발표 및 EU 개인정보보호기구 제동 사례, 2024.
- [36] Foundation Model Providers: Comparing OpenAI, Anthropic, Google and Others, AI Stack Authority, aistackauthority.com.
- [37] 개인정보보호법 제15조 - 개인정보의 수집·이용.
- [38] 개인정보보호법 제18조 - 개인정보의 목적 외 이용·제공 제한.
- [39] 개인정보보호법 제35조 - 열람요구.
- [40] 개인정보보호법 제36조 - 정정·삭제.
- [41] 개인정보보호법 제37조 - 처리정지.
- [42] 개인정보보호법 제28조의 8 - 국외이전 금지.
- [43] 개인정보보호법 제2조 - 정의.
- [44] 「생성형 인공지능의 저작물 학습에 대한 저작권법상 공정이용 안내서」, 문화체육관광부 & 한국저작권위원회, 2026.



참고문헌

- [45] 김병일, 인공지능 학습 데이터의 이용과 저작권 제한, 인공지능과 기본 권리에의 영향 제25-1호, 한국법제연구원, 2025.
- [46] 이인호, 인공지능과 개인정보보호의 긴장과 조화, 인공지능과 기본 권리에의 영향 제25-1호, 한국법제연구원, 2025.
- [47] (참고) 김세진, 호단 오마르, 통합적 접근의 명암: 한국 AI 기본법의 전략·진흥·규제 구조와 규제 리스크, ITIF, 2025.
- [48] 인공지능기본법, '고영향 AI' 개념 모호성 논란 <https://www.aitimes.com/news/articleView.html?idxno=168915>
- [49] 설명가능한 인공지능, https://spri.kr/posts/view/23296?data_page=1&code=industry_trend&study_type=&board_type=industry_trend&flg=0&page=5
- [50] 인공지능과 개인정보 보호의 긴장과 조화, <https://klri.re.kr/Platform/data/eaff417a-dee7-43dc-bfd3-696bf863c2b2/f626f6f7-eec5-400c-82dc-ae4f0c8e185c>
- [51] 한국정보통신기술협회(TTA) (2025). "신뢰할 수 있는 인공지능 개발 안내서"
- [52] 정란. (2025). 자동화된 의사결정 규제에 관한 비교법적 고찰 및 한국 법제에의 시사점.
- [53] 업무 효율을 높인 챗GPT 도입 사례(국내서비스에 해외모델이 사용되는 사례)<https://vizen.vizensoft.com//about/itinsight/read?no=48>
- [54] AI기본법 국내대리인 신고 현황, 엔트로픽·오픈AI·구글·메타 기업별 비교, 한국 데이터경제 신문 <https://www.dataeconomy.co.kr/news/articleView.html?idxno=40363>

감사합니다

한국의 AI·데이터 거버넌스는 주권적 관점에서 어떤 과제를 안고 있는가

국민주권적 AI 주권

정당성

실질 내용