

OECD 인공지능 글로벌 파트너십(GPAI) 제5차 총회 논의 동향

법무부 국제법무정책과 법무관 심 규 현

〈목 차〉

- | | |
|---|-----------------------|
| I. 회의 개요 | 7. OECD의 AI 관련 업무 현황 |
| II. 주요내용 | III. GPAI 워크숍 |
| 1. AI 정책 툴킷(AI Policy Toolkit) | 1. HAIP 보고 프레임워크 2.0 |
| 2. 에이전트형 AI(Agentic AI) | 2. 체화 AI(Embodied AI) |
| 3. 신뢰가능한 AI 투자(Trustworthy AI Investment) | IV. 향후 논의 방향 |
| 4. AI 컴퓨팅 | V. 관찰 및 평가 |
| 5. 각국의 AI 관련 정책 발표 | |
| 6. 관련 국제기구·협의체의 AI 관련 활동 공유 | |

I. 회의 개요

제5차 인공지능 글로벌 파트너십(Global Partnership on Artificial Intelligence, 이하 ‘GPAI’) 총회가 2026. 6. 9.부터 6. 11.까지 프랑스 파리 소재 OECD 본부에서 대면 및 비대면 병행 방식으로 개최되었다. 본 총회에는 48개 회원국을 비롯하여 시민정보사회자문위원회(Civil Society Information Society Advisory Council, 이하 ‘CSISEC’), 인터넷기술자문위원회(Internet Technical Advisory Committee, 이하 ‘ITAC’) 등 관련 기구 소속 대표단 약 300명이 참석하여, ‘신뢰할 수 있는 AI’를 목표로 하는 글로벌 AI 거버넌스 구축 방안에 관한 논의를 진행하였다.

금번 총회에서는 AI 정책 툴킷(AI Policy Toolkit)의 출시, 에이전트형 AI(Agentic AI) 연구, AI 투명성 프레임워크 등 그간의 성과가 소개되었다. 특히 논의 범위가 기존 OECD

AI 원칙의 이행 점점을 넘어 에이전트형 AI, AI 컴퓨팅, 데이터 거버넌스, 투자, 개방형 AI 생태계 등 다양한 의제로 확대되는 양상을 보였다. 그중에서도 에이전트형 AI는 실제 산업 현장으로 급속히 확산됨에 따라 인간 감독(Human Oversight) 및 책임성(Accountability) 확보를 위한 거버넌스 체계 구축이 핵심 정책 과제로 부상하였다.

II. 주요내용

1. AI 정책 툴킷(AI Policy Toolkit)

AI 정책 툴킷은 AI 정책 수립의 초기 단계에 있는 국가들도 활용할 수 있도록 사례 및 실무 중심의 구조로 설계되었다. 동 툴킷은 타국의 정책 사례를 열람할 수 있는 기능과 AI 기반의 정책 요약 기능을 제공하며, 요약과 함께 관련 문서 및 근거 문구를 병행하여 제시함으로써 투명성을 확보하고 있다.

특히 동 툴킷은 국가별 순위 산정이나 평가를 목적으로 하는 것이 아니라, 회원국의 정책 역량 강화와 국가 간 정책 학습의 촉진을 목적으로 설계되었다는 점이 강조되었다. 이러한 취지에 따라 설문 응답은 OECD에 저장되거나 공유되지 않는다. 사무국은 향후 동 남아시아 등으로 지역 협력을 확대할 계획이며, 특히 신뢰 가능하고 안전한 AI(Trustworthy, Safe and Secure AI)에 관한 정책 경험의 공유를 확대할 예정이라고 밝혔다.

2. 에이전트형 AI(Agentic AI)

OECD 사무국과 Tokyo Center of the GPA Expert Community는 공동으로 에이전트형 AI의 활용에 관한 연구 결과를 발표하였다. 관련 전문가 그룹에는 학계 및 산업계 등에서 약 30명의 전문가가 참여하고 있으며, 2026년 3월 개최된 제5차 AI WIPS(Work, Innovation, Productivity and Skills) 회의에서도 관련 연구 결과가 공유되어 80개국 이상에서 약 600명이 참석하는 등 높은 관심을 보인 바 있다.

발표에 따르면 에이전트형 AI의 도입은 빠르게 확산되고 있으며, 생산력 향상에 상당한 잠재력이 있는 것으로 평가된다. 다만 전문인력의 부족, 인간 감독 및 책임성 확보 등 거버넌스 체계 구축과 관련한 과제가 지속적으로 제기되고 있다는 점이 지적되었으며, 이에

따라 관련 거버넌스 문제의 해결이 최우선 과제로 제시되었다.

3. 신뢰가능한 AI 투자(Trustworthy AI Investment)

2025년을 기준으로 AI 관련 투자는 전 세계 벤처캐피털 투자의 61%를 차지하는 등 그 규모가 2022년 대비 약 2배 수준으로 크게 확대된 것으로 나타났다. 국가별로는 미국이 글로벌 AI 벤처투자의 과반을 점유하며 시장을 주도하고 있고, 중국이 그 뒤를 잇는 것으로 파악되었다.

4. AI 컴퓨팅

최근 2년간 채택 또는 갱신된 국가 AI 전략 중 70% 이상이 AI 컴퓨팅을 독립적인 핵심 정책 분야로 포함하고 있는 것으로 확인되었다.

이와 관련하여 대한민국은 AI 컴퓨팅 역량 확충을 위한 공급 측면의 지원 정책뿐만 아니라, 제한된 컴퓨팅 자원의 효율적 배분 및 활용을 위한 수요 측면 정책의 중요성도 함께 강조하였다. 프랑스는 지속가능한 AI 컴퓨팅을 위하여 탈탄소화된 에너지 공급이 중요하다는 점을 강조하였으며, 이스라엘은 AI 컴퓨팅을 국가 주권 확보를 위한 전략적 자산으로 인식하고 있다고 설명하였다.

5. 각국의 AI 관련 정책 발표

회원국들은 자국의 AI 관련 정책 추진 현황을 공유하였다. 대한민국은 APEC 의장국 활동과 연계한 AI 이니셔티브 및 아시아태평양 AI 센터 설립 추진 계획을 소개하였고, 스위스는 유럽평의회 AI 협약을 주요 기준으로 삼는 거버넌스 체계를, 포르투갈은 미성년자 보호를 위한 입법 검토 상황을, 이스라엘과 노르웨이는 각각 국가 차원의 AI 추진 체계 정비 및 공공부문 AI 도입 계획을 설명하였다. 국가별 발표 내용은 다음과 같다.

각국의 AI 관련 정책 발표	
대한민국	- APEC 의장국 활동의 일환으로 AI 주도 성장 및 역량 강화 등 AI 인프라 구축을 위한 APEC AI 이니셔티브를 추진 중이라고 소개함 - 아시아태평양 지역의 AI 역량 강화, 혁신 촉진 및 정보 공유를 위하여 아시아태평양 AI 센터의 설립을 추진 중이라고 설명함 - “AI for All” 비전 아래 유엔 등 국제기구와 협력하여 글로벌 AI 허브 구축 가능성을 검토하고 있으며, AI 인식 단계를 넘어 실질적인 AI 채택 및 성과 창출 단계를 지원하고자 한다고 밝힘
스위스	- EU AI 법을 직접 수용하지 않고 유럽평의회 AI 협약을 주요 기준으로 활용하고 있으며, 부문별 규제와 비규제 수단을 병행하는 AI 거버넌스 체계를 추진 중임 - 2027년 입법 초안 공개를 목표로 하고 있으며, 분야별 AI 행동규범과 윤리 가이드라인도 마련 중이라고 설명함
포르투갈	- 미성년자 보호를 위한 입법을 검토 중이며, 디지털 성년 기준을 16세로 설정하고 13세 미만 이용 제한 및 13~16세 이용 시 부모 동의 의무화를 논의 중임 - 관련 입법 과정에서 포르투갈 헌법 및 EU 기본권 헌장과의 충돌 가능성을 검토하고 있다고 설명함
이스라엘	- 국가 AI 국(National AI Directorate)을 신설하고 국가 차원의 AI 정책 추진 체계를 정비함 - 인프라 구축 및 사이버보안 등을 중심으로 하는 5개년 국가 AI 계획을 추진 중이라고 설명함
노르웨이	- 2030년까지 공공부문 전반에 AI를 도입한다는 목표를 제시하고, 이를 위한 국가 AI 연구 및 인프라 투자를 확대하고 있다고 설명함 - 정부 AI 전담 조직 신설 및 규제 샌드박스 예산 확보 등 AI 거버넌스 기반을 구축 중임

6. 관련 국제기구·협의체의 AI 관련 활동 공유

주요 국제기구 및 협의체는 각 기관에서 추진 중인 AI 관련 활동을 공유하였다. 전반적으로 AI 윤리 거버넌스, 국제표준의 개발, 역량 강화 등 각 기관의 전문 영역을 중심으로 활동이 전개되고 있는 것으로 확인되었으며, 그 구체적인 내용은 다음과 같다.

관련 국제기구·협의체의 AI 관련 활동	
UNESCO	- AI 윤리 거버넌스를 중심으로 분야별 AI 정책 및 역량 강화 활동을 추진
ITU	- AI 활용, 역량 강화 및 국제표준 개발 활동을 추진
세계은행	- AI 개발 및 거버넌스 지원을 위한 지식 생산과 국제협력을 추진
ISO/IEC	- 국제 AI 표준 개발 및 역량 강화 활동을 추진

7. OECD의 AI 관련 업무 현황

OECD 각 부서는 담당 분야에서 진행 중인 AI 관련 업무 현황을 발표하였다. 고용·노동 분야에서는 AI 노동시장 모니터링 체계의 구축을 위하여 15개국을 대상으로 근로자 및 고

용주의 AI 활용 실태에 관한 설문조사가 진행 중이며, 그 결과는 8월 중 데이터화될 예정이다. 반도체 분야에서는 국가별 반도체 산업 현황 보고서의 작성 및 비공식 정보 교류 네트워크가 운영되고 있으며, AI 데이터센터 내 반도체의 종류를 분류하고 공급망의 특정 국가 의존도를 분석하는 작업이 추진되고 있다.

스마트도시 분야에서는 도시의 AI 도입을 기술적 문제가 아니라 제도 및 운영 측면의 문제로 접근하여야 한다는 분석이 제시되었으며, 도시별 AI 준비 수준을 스스로 점검할 수 있는 진단 도구가 12월 중 발표될 예정이다. 보건 분야의 경우 OECD 회원국 중 별도의 보건 분야 AI 전략을 보유한 국가가 7개국에 불과한 상황으로, 관련 AI 실행계획이 2026년 가을에 발간될 예정이다.

III. GPAI 워크숍

1. HAIP 보고 프레임워크 2.0

본 세션에서는 히로시마 AI 프로세스(Hiroshima AI Process, 이하 ‘HAIP’) 보고 프레임워크 2.0의 개선 사항, 유럽평의회 AI 협약 등 타 거버넌스 체계와의 국제 정합성 및 기업의 이행 사례를 다루었다. HAIP 보고 프레임워크는 2023년 G7 디지털 장관회의를 계기로의 장국인 일본의 주도로 출범한 것으로서, 선진 AI 개발·배포 기업이 행동강령의 이행 등을 보고함으로써 관련 정보를 투명하게 공개하도록 하는 자발적 보고 제도이다.

사무국은 1차 보고 주기의 운영 결과 및 7개국 기업을 대상으로 진행한 시범 운영 과정에서 수렴한 의견을 반영하여 2.0 버전을 마련하였다고 설명하였다. 주요 개선 사항으로는 기업의 역할(모델 개발, 서비스 제공, 서비스 이용 등)에 따라 보고 항목을 구분한 점, 중소기업의 접근성을 높이기 위하여 응답 방식을 체크박스 형태로 간소화한 점, 에이전트형 AI 등 신기술 관련 항목을 추가한 점 등이 제시되었다.

이어 프랑스, 일본, 싱가포르 및 프랑스 AI 안전센터가 각각 자국 및 자사 차원의 이행 현황과 향후 협력 방향을 발표하였다. 프랑스는 G7 의장국으로서 HAIP 2.0에 56개 기업·기관의 서약 동참을 이끌어낸 점을 핵심 성과로 발표하면서, 중소기업부터 대기업까지 포용할 수 있는 프레임워크의 다양성과 거버넌스 간 상호운용성의 중요성을 강조하였다. 일

본은 68개 국가·지역 및 42개 기관이 참여하는 히로시마 프렌즈 그룹의 확대경과를 소개하고, HAIP가 글로벌 AI 거버넌스의 표준으로 발전할 수 있다고 평가하였다. 싱가포르의 자국의 AI 검증 프레임워크 발전 과정을 소개하고 2024년 초판 이후 2025년 생성형 AI까지 확장된 ASEAN AI 거버넌스·윤리 가이드를 설명하였으며, 현재 AI 검증, HAIP 행동강령 및 ASEAN 가이드 3자 간 매핑을 진행 중이라고 밝혔다. 프랑스 AI 안전센터는 HAIP 2.0과 유럽평의회 AI 협약을 비교 분석한 결과 약 3분의 2의 정합성을 보였으며, 양자가 상호보완적 관계에 있다고 분석하였다.

기업 패널 세션에서는 마이크로소프트, 프랑스의 법률 AI 스타트업(Leaid), 안전성 평가 전문 기업(Safer AI) 및 인도 IT기업(Infosys)이 HAIP 이행 경험을 공유하였다. 마이크로소프트는 HAIP 제출 보고서의 작성 과정에서 모델 개발자, 서비스 제공자 및 배포자 사이의 리스크 커뮤니케이션이 개선되는 효과를 확인하였다고 밝히며, 자사의 안전성 점검 체계 및 모의 공격 테스트 운영 사례를 공유하였다. Leaid는 중소기업의 관점에서 HAIP가 국제 시장 진출 시 준거할 수 있는 명확한 규범 기반을 제공한다고 평가하면서, 자율 AI 시스템에 대한 인간의 감독을 어떻게 유지할 것인가 향후 핵심 과제라고 언급하였다. Safer AI는 HAIP 2.0에서 기업의 역할을 구분하고 체크박스 방식을 도입한 점이 중소기업의 접근성을 높였다고 평가하면서도, AI 역량이 빠르게 발전함에 따라 안전성 평가 결과의 일부가 기밀로 처리되는 경향이 투명성을 저해할 수 있다는 우려를 제기하였다. Infosys는 위험관리, 투명성, 연구개발 등을 포괄하는 종합적인 단일 보고 프레임워크의 공백을 HAIP가 메우고 있다고 평가하며, 향후 중소기업 및 신흥국 기업의 참여를 확대하고 연간 보고 분석을 강화할 필요가 있다고 제안하였다.

2. 체화 AI(Embodied AI)

OECD와 글로벌 기술 포럼이 공동으로 기획한 본 세션에서는 체화 AI의 기술 현황 및 정책적 합의 등에 관한 발표가 이루어졌다. 여기서 체화 AI란 로봇, 자율주행차 등 물리적 장치에 탑재되어 주변 환경을 인식하고 직접 행동하는 형태의 AI를 지칭한다.

기조 발표를 맡은 MIT 컴퓨터과학·로봇공학과의 Leslie Kaelbling 교수는 주변 상황을 실시간으로 파악하고 이에 따라 동작을 수행하는 학습 방식인 체화 AI의 개념과 자율주행차, 드론 군집 등의 활용 예시를 설명하였다. Kaelbling 교수는 체화 AI가 데이터 수집 비

용 및 시뮬레이션의 한계 등으로 인하여 LLM(대규모 언어 모델, Large Language Model)에 비하여 발전 속도가 느리며, 단기적으로는 물류창고 자동화 및 공장 운영 등에서 우선 활용될 것이나 범용 가사 로봇의 상용화는 아직 요원하다고 전망하였다. 아울러 스타트업의 데모 시연을 과신하지 않는 신중한 태도가 필요하며, 군사적 활용에 대한 경계가 필요하다는 점을 강조하였다.

구글은 카메라 및 센서로 주변을 인식하는 인공지능과 실제 동작을 수행하는 인공지능을 결합한 로봇 기술(Gemini Robotics)을 소개하면서, 로봇이 사람이라면 당연히 아는 기초적인 물리 상식을 학습할 데이터가 부족한 점이 가장 큰 과제라고 분석하였다. 또한 시각 장애인 지원, 농업 지원, 공장 신입 직원 교육 지원 등의 응용 사례를 소개하는 한편, CCTV 영상 분석을 통한 보안 체계 역추적 등 체화 AI의 악용 위험성을 지적하며 이와 관련하여 영국 당국과 협력을 진행 중이라고 밝혔다. 프랑스 원자력·재생에너지청은 체화 AI를 성공적으로 도입하기 위해서는 기술력보다 제도적 기반과 사람-AI 간 협업 방식의 설계가 더 중요하다고 강조하고, 실제 현장 적용을 위한 신뢰성 검증·인증 체계의 마련이 시급하다고 제언하였다.

질의 및 토론에서 미국은 데모 환경에서 검증된 성능이 실세계로 일반화되는지를 검증하는 방법에 관하여 질의하였고, 발표자는 하드웨어 자체의 성능은 신뢰할 수 있으나 AI의 일반화 능력에 대해서는 과도한 신뢰를 경계하여야 한다고 답변하였다. EU는 생성형 AI와 로봇공학의 결합에 있어 가장 큰 기술적 장애물이 무엇인지 질의하였고, 발표자는 3차원 공간에 대한 이해 및 행동의 인과관계를 모델링하는 능력의 부재를 핵심 장애물로 제시하였다. 영국은 인간의 안전을 보장할 수 있는 수준과 그 인증 방법론에 관하여 질의하였으며, 발표자는 힘의 제한에서 작업 범위의 한정을 거쳐 성능 요건의 충족에 이르는 순서로 단계적으로 인증 체계를 구축해 나가는 방식을 제안하였다. 이 밖에 다수의 전문가가 연구 분야의 다양성 확보가 중요하다는 데 공감하면서, 특정 접근법으로 연구가 획일화되는 경향에 대한 우려를 공유하였다.

IV. 향후 논의 방향

사무국은 GPAI의 향후 업무 방향을 개념 정립, 증거 기반 강화, 분야별 AI 도입 및 기술

발전 추적의 4개 축으로 구성하여 제시하였다. 개념 정립은 AI 스택, AI 개방성, 에이전트형 AI 등 핵심 개념의 정의 및 분석 틀을 정립하는 작업이고, 증거 기반 강화는 OECD AI 정책 관측소 등을 통하여 국가별·지역별 AI 정책 데이터 및 지표를 확충하는 작업이다. 분야별 AI 도입은 농업, 보건, 노동시장 등 부문별 AI 활용 현황 및 도입 장벽을 분석하는 작업이며, 기술 발전 추적은 체화 AI, 에이전트형 AI 등 신기술 동향에 대한 선제적 모니터링을 의미한다.

이에 대하여 영국은 개념 정립을 GPAI 업무 중 가장 부가가치가 높은 영역으로 평가하면서, OECD 내 다른 위원회의 AI 관련 업무와의 중복을 방지하기 위하여 업무 분야를 명시적으로 정리할 필요가 있다고 제안하였다. 네덜란드는 농식품 분야에서 AI를 안전하게 활용하기 위해서는 상호운용 가능한 데이터 접근, 맞춤형 설계, 견고한 인프라 거버넌스 및 사이버보안 대책이 필요하다고 제안하였다. 덴마크는 중점 업무에 있어 선택과 집중이 필요하다고 강조하며, AI 기술구조의 기반이 되는 분야(컴퓨팅, 데이터 거버넌스 등)에 대한 심화 분석과 정부의 AI 활용 현황에 대한 연구 확대를 제안하였다.

캐나다는 공공 신뢰, 경제성장 및 디지털 주권의 3대 축으로 구성된 자국 AI 전략과 GPAI 업무 간의 연계성을 확인하면서, AI 도입을 가로막는 장벽에 대한 분석의 강화를 요청하였다. 브라질은 개념 정립 작업이 실무 가이드 및 정책 옵션 등으로 구체화되어야 실질적 효용이 있다고 강조하며, 농업·보건·교육·기후 등 공공 이익 분야의 실제 도입 문제를 다루고 신기술 추적 작업 역시 책임성, 안전, 프라이버시 및 노동시장 영향까지 포괄하여야 한다고 제안하였다. 이집트는 AI 조달, 규제 샌드박스 및 아동 보호 가이드라인 마련의 중요성을 강조하였고, 다음 달 제네바에서 개최되는 UN 차원의 AI 대화와 GPAI 업무 간 연계의 필요성을 언급하였다. BIAC는 리스크 기반·비례적·상호운용 가능한 거버넌스 접근의 유지를 강조하면서, AI로 인한 노동시장 변화에 대한 모니터링을 지속하여 줄 것을 요청하였다.

V. 관찰 및 평가

금번 총회는 기존의 논의 의제를 넘어 새로운 정책 의제로 논의의 범위가 확대되는 양상을 보였다. 특히 AI 컴퓨팅 자원의 확보, AI 투자의 확대, 국가 차원의 AI 인프라 구축

등이 주요 정책 의제로 논의된 점은, 글로벌 AI 경쟁이 기술 개발의 차원을 넘어 컴퓨팅 역량 및 투자 유치 등 국가 경쟁력 확보의 차원으로 확대되고 있음을 시사한다.

회원국들은 국가별 여건에 따라 다양한 AI 정책 및 규제 체계를 추진하고 있으나, AI 역량 강화, 공공부문에서의 활용 확대, 안전하고 신뢰할 수 있는 AI 개발 및 국제협력 강화라는 공통된 정책 방향을 공유하고 있는 것으로 평가된다. 아울러 UNESCO, ITU, 세계은행, ISO/IEC 등 관련 국제기구들도 AI 윤리, 국제표준 및 역량 강화 등의 분야에서 활동을 확대하고 있어, 향후 GPAI와의 연계 및 협력이 더욱 강화될 것으로 전망된다.

대한민국은 APEC AI 이니셔티브 및 아시아태평양 AI 센터 설립 등 국제협력 의제를 적극적으로 제안하였는바, 향후 GPAI에서 AI 역량 강화 및 글로벌 AI 거버넌스 협력 분야 등에서 주도적인 역할을 수행할 수 있을 것으로 기대된다. 한편 GPAI 회원 확대와 관련하여서는 수적 확대보다 지역 대표성 및 기여 가능성을 중심으로 하는 전략적 접근으로 방향을 정립할 필요성이 제기된다.

끝으로 체화 AI에 관한 논의에서는 기술 발전 속도에 비하여 거버넌스의 공백이 두드러진다는 점이 확인되었다. 따라서 일반화 능력의 검증, 군사적 응용 및 물리적 안전 인증 등 정책적 과제에 대한 선제적 검토가 요구되며, 우리나라 역시 이러한 논의 동향을 지속적으로 주시하면서 국내 AI 관련 법제 및 정책에 반영할 실무적 필요성이 있다고 판단된다.