
인공지능 활용 원칙 (AI Principles)

1. OECD, 인공지능 활용 원칙 권고안 채택

■ (개요) 미국 등 42개국은* OECD 각료이사회 회의를 통해 '인공지능 활용 원칙에 관한 이사회 권고안'을 공식 채택('19.5월)

* OECD 36개 회원국 + 브라질 등 6개국

- 권고안은 AI 시스템을 건전하고, 안전하고, 공정하고, 신뢰할 수 있게 디자인해야 한다는 기본 이념하에,
- AI와 관련하여 모든 사람이 공유해야 하는 '①가치관에 대한 일반 원칙'과 '②회원국 정부를 향한 정책 권고 사항'으로 구성

■ (중요성) 특히 해당 권고안은 AI 기술 윤리 등에 관해 국제사회가 공동으로 내놓은 첫 번째 원칙¹⁾이라는 점에서 주목됨

- OECD에서 제시한 AI 원칙은 향후 AI 활용 및 윤리지침에 대한 국제표준으로 참고될 가능성이 매우 높다는 점에서 중요²⁾
- 실제 '19.6월 오사카에서 개최된 G20 디지털 장관회의의 선언문에서는 OECD의 권고안 내용을 거의 그대로 인용한 'G20 인간 중심의 AI 원칙'이 공식 채택됨

1) 이전에도 일본, EU 등에서 개별적으로 AI 윤리를 규정하거나 기업에서 자체적인 AI 윤리 원칙을 내놓은 적은 있음

2) 2013년 OECD가 결의한 '개인정보 보호지침'은 미국, 유럽 등 전 세계 '개인정보 보호법'의 기초가 되었음

■ (내용) AI 권고안에 포함된 ‘일반 원칙’과 ‘정책 권고 사항’은 각각 다음의 5가지 항목으로 이루어져 있음

(1) 책임 있는 AI 시스템 관리를 위한 일반 원칙

[표1] OECD AI 일반 원칙의 5가지 항목

항 목	주요 내용
① 포용적 성장, 지속 가능한 개발과 행복 증진	· AI 활용 시 인류의 능력과 창의성을 증대시키고, 여러 형태의 불평등(성, 지역 등)을 줄여나가며, 자연을 보호함으로써 포용적 성장, 지속 가능한 개발 및 행복 증진을 추구해야 한다.
② 인간 중심적 가치와 공정성	· AI 시스템은 법의 권위, 인권, 민주적 가치관, 다양성을 존중하도록 설계되어야 한다. · AI 시스템은 공정한 사회를 보장하기 위해 적절한 보호 장치를 갖추어야 한다.
③ 투명성과 설명 가능성	· 사람들이 AI에 기반을 둔 결과를 이해하고 이를 평가할 수 있도록 유의미한 정보를 제공함으로써 AI 시스템에 관한 투명성과 책임 있는 설명을 보장해야 한다.
④ (시스템의) 안전성, 안정성 및 견고성	· AI 시스템은 안전하고 안정적으로 작동하여 상황과 여건에 관계없이 비합리적인 위험을 초래하지 않아야 한다. · AI 시스템은 시스템 내의 데이터, 프로세스, 의사결정과 관련한 ‘추적 가능성’을 갖추어야 하며, 시스템을 통해 산출된 결과를 분석할 수 있어야 한다. · 개인정보, 디지털 보안, 편향적 의사결정 등과 같은 시스템 관련 위험을 대처하기 위해 체계적인 리스크 관리 체계를 갖추어야 한다.
⑤ 책임성	· AI 기술을 활용하는 개인 및 조직은 위 4가지 원칙에 따라 시스템을 적절하게 운용할 책임이 있다.

*출처: OECD 홈페이지

(2) 국가 정책 권고 사항

[표2] OECD AI 정책 권고 사항의 5가지 항목

항 목	주요 내용
① AI 연구 개발에 대한 투자	<ul style="list-style-type: none"> · 정부는 AI 기술의 혁신을 이끌어내기 위해 공공 및 민간 연구개발 투자를 촉진해야 한다. · AI 연구개발 환경을 지원하고 'AI 표준'의 사용과 상호 운용성을 증진하기 위해 개인정보, 데이터 보호에 힘써야 한다.
② 디지털 생태계 조성	<ul style="list-style-type: none"> · 디지털 기술, 인프라, AI 지식을 공유하는 디지털 생태계의 개발을 촉진하고 접근성을 향상시켜야 한다. · 이를 위하여 정부는 데이터의 안전하고, 공정하고, 합법적이며, 윤리적인 공유를 지원하는 장치를 마련해야 한다.
③ 유연한 정책 환경	<ul style="list-style-type: none"> · 정부는 AI 시스템의 연구 개발 단계에서 실제 적용으로의 민첩한 전환을 지원하는 정책 환경을 만들어야 한다. · AI 시스템의 혁신과 건전한 경쟁을 장려하기 위해 관련 정책과 규제, 시스템 평가 체계를 적절하게 검토하고 조정해야 한다.
④ 인적 역량 배양 및 일자리 변혁 대응	<ul style="list-style-type: none"> · 정부는 사람들에게 AI 기술을 습득하게 하여 AI 시스템을 효과적으로 사용할 수 있도록 도우며, 변화된 노동시장에서 근로자가 새로운 기회를 찾을 수 있도록 지원한다. · AI 기술의 활용이 일자리의 질과 안정성을 높이고, 생산성에 도움을 주며, AI 기술의 활용으로 인한 이익이 폭넓고, 공평하게 공유될 수 있도록 지원한다.
⑤ 국제협력	<ul style="list-style-type: none"> · AI 관련 지식을 공유하고, 표준을 개발하며, AI 관리·감독의 책임을 다하기 위해 국경과 산업 분야를 넘어서 협력한다.

*출처: OECD 홈페이지

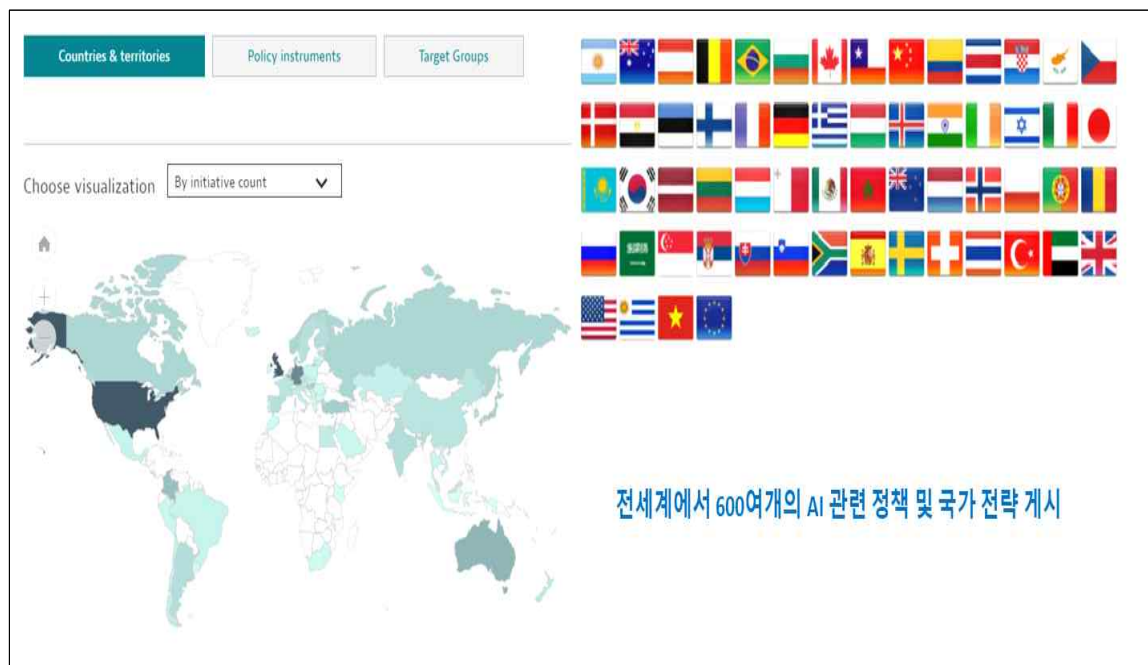
■ (정책관측소) '20.2월 온라인 AI 정책관측소를* 설립하고 각국의 AI 관련 정책 정보를 수집

* AI Policy Observatory (홈페이지: OECD.AI)

- OECD 회원국이 AI 권고안을 이행하는지 여부, 즉, 신뢰할 수 있고 사회에 유익을 주는 AI 시스템을 개발하는지를 지속적으로 모니터링
- AI 정책에 대한 국가 간 대화를 촉진하고, 모범 사례를 공유하는 허브를 구성하는 것을 목표로 함
- (블로그) 또한, 정책관측소 내에서 AI 블로그*를 운영하며 AI 전문가들의 연구 결과와 시스템 적용 등에 대한 경험을 공유하는 공간으로 활용 중

* AI Wonk blog

[그림1] 국가별 AI 관련 정책 분포도



*출처: OECD AI 정책관측소 홈페이지

2. 미국 NAIC, 인공지능(AI) 지도 원칙을 만장일치로 채택

■ (개요) NAIC 회원들은 인공지능 지도 원칙을 만장일치로 채택(20.8월)

- 이 원칙은 AI 기술을 보험업에 적용할 시 요구되는 일반적 사항들에 대한 정보를 제공하기 위해 마련됨
 - 단, 강제적인 법적 효력은 없으며 보험회사, 전문가, 감독·정책당국 및 기타 이해관계자 등의 AI 관련 담당자*들을 지원하기 위한 지침으로 개발됨

* 이하 'AI 활용자'

- NAIC 산하의 '빅데이터 & AI 워킹그룹'*에서 해당 원칙을 개발하였으며, OECD의 AI 원칙³⁾에 기반을 두고 있음

* Big Data and Artificial Intelligence Working Group

■ (배경) 빅데이터, AI 등 새로운 디지털 기술의 접목으로 보험산업의 구조 변화가 급속히 진전되는 상황

- 마케팅, 언더라이팅, 고객관리, 보험금 지급, 보험사기 방지 등 보험업 가치 사슬(Value-Chain)의 전(全) 단계에서 AI 기술이 활용되고 있음
- 감독·정책당국은 AI 기술을 통한 디지털 혁신을 지원하는 동시에 보험시장 안정, 소비자·투자자 보호에도 적합한 규제체제 마련 필요성 증대

- 특히, 관련 기술이 야기할 수 있는 윤리적 문제* 등에 대한 선제적 논의가 필요

* AI 는 의사결정 과정 및 결과를 명확하게 이해하기 어려운 'Black Box' 의 속성을 갖고 있음

3) 경제협력개발기구(OECD)에서 2019년 5월 발표한 AI 원칙이며, 미국을 포함한 42개국이 채택

■ (원칙) NAIC의 AI 원칙은 다섯 가지 주요 원리를 포함(FACTS)⁴⁾

- ① (공정성 및 윤리성) AI 시스템은 공정하고 윤리적으로 개발되고 활용되어야 함
 - AI 기술을 활용할 시, 특정 집단(예: 인종, 성별 등)에 대한 차별을 금지
 - AI 시스템은 사람들에게 피해를 주거나 의도치 않은 결과가 발생하지 않도록 실행되어야 하며, 만약 그런 일이 발생할 경우 시스템을 적절히 수정해야 함
- ② (책임성) AI 활용자는 AI 시스템 운영에 대한 책임감을 가지고 있어야 함
 - AI 시스템을 통해 산출된 최종 결과값의 기반이 되는 데이터는 관련 보험법과 규정에 따라 보관(retained)되어야 하고 생성할 수 있어야 함
 - AI 활용자는 AI 시스템의 내재적인 리스크 특성을 고려하여 법과 규정에 부합하게 운영되도록 안전장치를 마련해야 하며 지속적인 모니터링을 해야함
- ③ (법규 준수) AI 활용자는 적용 가능한 모든 보험법과 규정을 준수하기 위해 필요한 지식과 자원을 마련해야 함
 - AI 시스템이 속해 있는 지역(주)에 의해 우선적으로 규제를 적용받고, 각 관할권의 보험법과 규정은 반드시 준수해야 한다는 것을 인지해야 함
- ④ (투명성) 공공의 신뢰를 높이기 위해 AI 시스템에 관한 적극적인 공개를 약속
 - AI 활용자는 관련 이해관계자에게 AI 시스템에 사용된 데이터의 종류, 사용한 목적, 그리고 최종 결과에 대해 공개
 - 법과 규정에 따라, 이해관계자는 AI가 주도하여 내린 의사결정에 대해 문의 및 검토할 수 있는 방안이 있어야 하며, 해당 문의를 통해 얻은 정보는 이해하기 쉽고 의사결정에 영향을 미친 요인들을 설명할 수 있어야 함

4) **FACTS** : Fair and Ethical / **A**ccountable / **C**ompliant / **T**ransparent / **S**ecure, Safe and Robust

⑤ (안전성 등) AI 시스템은 상황과 여건에 관계없이, 법과 규정에 적합하도록 안전하고 안정적으로 작동해야 함

- 이를 위해 AI 활용자는 시스템 운영 시 사용되는 데이터, 프로세스 및 의사 결정과 관련한 합리적인 수준의 추적 가능성을 확보해야 하며, 시스템을 통해 산출된 결과물을 분석할 수 있어야 함
- 법과 규정에 정의된 개인정보, 디지털 보안, 부당한 차별 금지 등과 관련된 위험을 대처하기 위해 '체계적인 리스크관리' 적용 필요

