KISDI Perspectives

동향 September 2024 No.3

인공지능과 인권, 민주주의, 법치주의에 관한 유럽평의회 프레임워크 협약 현황

이나경 연구원
정보통신정책연구원 디지털국제협력연구센터

KISDI 정보통신정책연구원
KOREA INFORMATION SOCIETY DEVELOPMENT INSTITUTE

KISDI Perspectives

동향

2024. 9. 25

인공지능과 인권, 민주주의, 법치주의에 관한 유럽평의회 프레임워크 협약 현황

이나경 연구원
정보통신정책연구원 디지털국제협력연구센터, nlee@kisdi.re.kr

요 약

- 2024년 5월, 유럽 평의회(Council of Europe)는 인공지능이 인권, 민주주의, 법치주의에 미치는 악영향을 완화하기 위해 세계 최초의 법적 구속력을 가진 '인공지능과 인권, 민주주의, 법치주의에 관한 유럽 평의회 프레임워크 협약*'을 채택함
  * "Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law"
- (주요 내용) 협약은 인공지능 시스템의 설계, 개발, 사용, 폐기에 이르는 전 생애주기 동안 인권, 민주주의, 법치주의를 보호하고 기술 혁신 촉진을 주요 목표로 함
  - 이를 위해 △투명성, △책임성, △평등 및 차별 금지, △장애인 및 아동의 권리 보호, △디지털 리터러시 역량 강화, △개인정보 보호, △신뢰성 등의 원칙을 제시하며, 각 원칙이 인공지능 시스템의 전과정에서 준수되도록 요구함
  - 적용 대상은 인권, 민주주의, 법치주의를 침해할 가능성이 있는 공공 당국과 이를 대신하는 민간 행위자이며, △국가 안보 이익 보호 및 국방과 관련된 활동, △아직 사용되지 않은 인공지능 시스템에 관한 연구 및 개발 활동은 제외됨
- (협약의 의의) 본 협약은 세계 최초의 법적 구속력을 갖춘 AI 협약으로서, 유럽 평의회 비회원국과 비국가 행위자들의 가입과 참여도 권장하고 있어 글로벌 표준으로 자리 잡을 가능성이 있음
- (우리나라 동향 및 제언) 본 협약 가입은 우리나라의 인공지능 분야 국제협력을 강화하고, 인공지능 기본법 제정의 동력을 얻어, 인권, 민주주의, 법치주의를 보호하는 법적 기반을 마련할 계기가 될 수 있음
  - 우리나라에서도 인공지능 기술로 인한 인권침해 및 민주주의 훼손 사례가 발생하고 있어 이를 방지하기 위한 여러 법안과 가이드라인이 마련되었으나, 이를 포괄하는 인공지능 기본법은 부재한 상황
  - 우리나라는 AI 서울 정상회의('24년 5월)에서 '서울 선언'을 채택하며 AI 분야에서 민주주의적 가치, 법치주의 및 인권·기본권 자유와 프라이버시를 보호하고 증진하겠다는 의지를 밝힘

01 개요 및 배경

- 인공지능 기술의 발전은 인간과 사회의 웰빙, 지속 가능한 개발, 성평등을 증진할 수 있는 잠재력이 있으나, 동시에 인간의 존엄성과 개인의 자율성, 인권, 민주주의, 법치주의를 훼손하는데 사용될 수 있음
  - 인공지능 기술은 허위 정보를 생성하거나 취약 계층의 민주적 절차 참여를 배제하는 등 민주주의와 인권에 위협을 가하고, 공적 영역을 분열하고 시민들의 신뢰를 약화시킬 수 있음
- 2024년 5월 17일, 유럽 평의회(CoE, Council of Europe)*는 세계 최초로 국제적으로 법적 구속력이 있는 조약인 "인공지능과 인권, 민주주의, 법치주의에 관한 유럽평의회 프레임워크 협약(이하 협약, Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law)"을 채택함
  * 유럽 지역의 선도적인 인권 보호 기관으로, 1949년에 설립된 이래로 46개 회원국에 걸쳐 유럽 인권 협약(ECHR)을 중심으로한 공동 법적 체계를 구축[1]
- 유럽 평의회는 인공지능 시스템 생애주기 내 활동의 혜택을 극대화하면서도, 인권, 민주주의, 법치주의에 미치는 위협을 완화하기 위해, 이러한 활동을 관리하는 전 세계적으로 적용가능한 법적 프레임워크의 필요성을 인식하여 본 협약을 수립함
  - 본 협약은 인공지능 위원회(CAI)의 2년간 협의를 통해 완성되었으며, 유럽 평의회 46개 회원국, EU, 11개의 비회원국*이 협약을 작성하고, 민간 부문, 시민 사회 및 학계 대표자들이 참관인으로 협상에 참여함
    * 아르헨티나, 호주, 캐나다, 코스타리카, 바티칸, 이스라엘, 일본, 멕시코, 페루, 미국, 우루과이
  - 본 협약은 EU 인공지능 법안과 유사한 위험 기반 접근 방식을 채택하여, 인공지능 시스템의 설계, 개발, 사용, 폐기에 이르는 생애주기 전반에서 인권, 민주주의, 법치주의를 준수하며 기술 진보와 혁신을 촉진함[2]
- 본 협약은 2024년 9월 5일 리투아니아 빌뉴스에서 열리는 법무장관 회의에서 유럽 평의회 회원국 및 비회원국 모두에게 개방될 예정이며, 주요 추진 경과는 다음과 같음[3];

1 Council of Europe. "Key Facts," https://www.coe.int/en/web/portal/the-council-of-europe-key-facts.

2 dig.watch (2024.05.20). "Council of Europe adopts first-ever international treaty on AI", https://dig.watch/updates/council-of-europe-adopts-first-ever-international-treaty-on-ai.

3 Council of Europe, "Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law brochure."; dig.watch. "Convention on AI and human rights (CoE process)", https://dig.watch/processes/convention-on-ai-and-human-rights-council-of-europe-process.

- (´19년) 인공지능에 관한 임시 위원회(CAHAI)가 협약의 실행 가능성을 검토하기 위한 작업 시작
- (´21년) 유럽 평의회의 장관위원회에서 인공지능 시스템의 개발, 설계, 적응에 관한 법적 장치를 마련하는 임무를 수행할 인공지능 위원회(CAI)의 설립을 승인
- (´22년) 인공지능에 관한 임시 위원회(CAHAI)는 인공지능 위원회(CAI)로 승격되었으며, CAI는 협약의 본문을 작성하고 협상 시작
- (´24년 3월) 인공지능과 인권, 민주주의, 법치주의에 관한 프레임워크 협약의 초안 최종 확정
- (´24년 5월 17일) 인공지능과 인권, 민주주의, 법치주의에 관한 프레임워크 협약 채택
- (´24년 9월 5일) 협약 개방

● (구성) △1장 일반규정, △2장 일반 의무, △3장 인공지능 시스템의 생애주기 관련 원칙, △4장 구제책, △5장 위험과 부정적 영향 평가 및 완화, △6장 협약의 이행, △7장 후속 메커니즘 및 협력, △8장 최종 조항

● 본고는 '인공지능과 인권, 민주주의, 법치주의에 관한 유럽 평의회 프레임워크 협약'의 주요 내용을 소개하고자 함

02 인공지능과 인권, 민주주의 및 법치주의에 관한 유럽 평의회 프레임워크 협약의 주요 내용[4]

▌제1장. 일반규정 : △목적과 취지, △인공지능 시스템의 정의, △적용 범위

● (목적과 취지) AI 시스템의 생애주기 전반에 걸친 활동이 인권[5]*, 민주주의[6]**, 법치주의[7]*** 원칙의 준수를 보장하고, 기술 진보와 혁신을 촉진

4 Council of Europe. "Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law." Council of Europe Treaty Series 225, Council of Europe, 2024.; Council of Europe.; "Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law." Council of Europe Treaty Series 225, Council of Europe, 2024.

5 Council of Europe. "What Are Human Rights?", https://www.coe.int/en/web/portal/what-are-human-rights.

6 Council of Europe . "Democracy.", https://www.coe.int/en/web/compass/democracy#What%20is%20Democracy.

7 Council of Europe, Venice Commission (2011). "Report on the Rule of Law (CDL-AD(2011)003rev).", https://www.venice.coe.int/webforms/documents/?pdf=CDL-AD(2011)003rev-e.

* 인권은 인간을 개별 존재와 인류의 일원으로서 존중하고 인간의 존엄성을 보장하는 것

** 민주주의는 국민의 의지에 따라 통치하는 '국민의 힘(power of the people)'으로 볼 수 있으며, 이론적으로 모든 국민의 "의지"에 따라 국민을 위한 정부를 의미함

*** 법치주의는 국가 내 모든 개인과 권위자가 공적이든 사적이든 상관없이 공공적으로 제정된 법에 구속되고, 그 법이 일반적으로 미래에 효력을 발휘하여 법원에서 공정하게 집행되는 혜택을 누릴 권리를 가지는 것

- (인공지능 시스템의 정의) 명시적 혹은 암시적 목표를 달성하기 위해 입력된 정보로부터 물리적 또는 가상 환경에 영향을 미치는 예측, 콘텐츠 생성, 추천, 의사 결정을 수행하는 기계 기반 시스템
  - 본 정의는 OECD의 AI 권고안(OECD/LEGAL/0449, 2019, 2023 개정)을 참고하여 작성되었으며, 국제기구인 OECD의 정의를 사용함으로써 인공지능 분야 국제협력을 강화하고, 각 당사국의 국내 법체계 내에서 인공지능과 관련된 다양한 법적 수단의 일관된 적용을 촉진할 수 있음
- (적용 범위) 인권, 민주주의, 법치주의를 침해할 가능성이 있는 공공 당국 또는 이를 대신하는 민간 행위자의 인공지능 시스템의 생애주기 내 활동
  - (예외) 국가 안보 이익 보호와 관련된 인공지능 시스템의 생애주기 내 활동*, 국방 관련 사항**, 아직 사용되지 않은 인공지능 시스템에 관한 연구 및 개발 활동

* 적용 가능한 국제법(예: 국제인권법)에 부합하는 방식으로 수행되어야 함

** 국방과 관련된 인공지능 시스템의 생애주기 내의 활동이 국제법의 적용을 받지 않는다는 의미는 아님

제2장. 일반 의무: △인권 보호, △민주적 절차의 완전성과 법치주의 존중

- (인권 보호) 인공지능 시스템의 생애주기 내 활동이 관련 국제법 및 국내법에 명시된 인권 보호 의무와 일치하도록 보장하는 조치를 채택·유지해야 함
- (민주적 절차의 완전성과 법치주의 존중) 인공지능 시스템이 민주적 제도의 무결성, 독립성, 효율성을 훼손하지 않도록 보장하고, 개인의 자유로운 의견 형성 및 공개 토론 참여 등 민주적 절차를 보호하는 조치를 채택·유지해야 함
  - 선거 과정에서 악의적인 외국의 간섭에 대한 사이버 보안 조치 혹은 잘못된 정보·허위 정보 확산을 대응하기 위한 조치 등을 포함함

제3장. 인공 지능 시스템의 생애주기 내 활동과 관련된 원칙: △일반적 접근 방식, △인간의 존엄성과 개인의 자율성, △투명성과 감독, △책임성과 책무, △평등과 차별 금지, △프라이버시 및 개인정보 보호, △신뢰성, △안전한 혁신

- 인간의 존엄성과 개인의 자율성 존중
  - (인간의 존엄성 존중) 배경, 특성, 환경에 관계없이 각 개인의 고유한 가치를 존중해야 함
  - (개인의 자율성 존중) 인공지능 기술이 개인의 삶에 미치는 영향을 통제하여 인공지능으로 인해 주체성과 자율성이 약화되지 않도록 해야 함
- (투명성과 감독) 인공지능 시스템이 생성한 콘텐츠를 식별*할 수 있도록 투명성과 감독 요구사항을 보장하는 조치**를 마련하고 지속적으로 유지해야 함

  * 인공지능 시스템의 출력물에 인식 가능한 서명을 삽입하는 라벨링 및 워터마킹과 같은 기술이 포함되며 기술의 가용성과 효과가 입증된 경우에만 사용 가능

  ** (예시) 데이터 출처, 훈련 방법론, 데이터 소스의 유효성, 사용된 훈련, 테스트 및 검증 데이터에 대한 문서화 및 투명성, 위험 완화 노력, 이행된 프로세스 및 결정과 같은 주요 고려사항 기록

  - (투명성) 인공지능 생애주기 내 활동 거버넌스의 개방성과 명확성을 의미하며, 투명성을 위한 정보 공개는 프라이버시, 기밀, 국가 안보, 제3자의 권리 보호, 공공 질서, 사법 독립 등과 충돌할 수 있으므로, 당사국은 이러한 상충하는 이익 간 균형을 유지하고 인권법의 기본 체계를 변경하지 않고 조정해야 함
  - (감독) 인공지능 시스템의 생애주기 내 활동을 모니터링, 평가 및 안내하기 위해 고안된 메커니즘, 프로세스 및 프레임워크*를 의미하며 내부 고발자 보호도 옵션이 될 수 있음

    * (예시) 법률, 정책 및 규제 프레임워크, 권장 사항, 윤리 지침, 실행 강령, 감사 및 인증 프로그램과 당국(감독 기관 및 위원회, 데이터 보호 기관, 평등 및 인권 기관, 국가인권기구 혹은 소비자 보호 기관 등)의 모니터링, 교육 및 인식 프로그램 등
- (책임과 의무) 당사국은 인공지능 시스템의 생애주기 내 활동에 책임이 있는 개인이나 조직이 인권, 민주주의, 법치주의에 부정적 영향을 미칠 경우, 이를 책임질 수 있는 메커니즘*을 마련해야 함

  * 사법 및 행정 조치, 민사, 형사 및 기타 책임 제도, 공공 부문의 경우 결정에 이의를 제기할 수 있는 행정 및 기타 절차 또는 운영자에게 특정 책임과 의무를 부과하는 것도 포함됨
- (평등과 차별 금지) 국내외 인권 의무에 따라 인공지능 시스템의 생애주기 전반에서 불평등과 차별 금지를 보장하는 조치를 채택·유지해야 함
  - 본 협약은 인공지능 시스템의 생애주기 내 활동으로 인해 영향을 받는 당사자의 의견을 반영할 것을 권고함

- 본 협약은 인공지능 시스템의 생애주기 내 활동에서 발생할 수 있는 편향(bias)을 세분화하고 편향으로 인한 부정적 영향을 방지하기 위해 이를 반영하도록 권장하며, 편향의 종류는 다음과 같음;

편향의 종류

잠재적 편향	알고리즘 개발자의 무의식적·의식적인 고정관념이나 편견, 시스템이 구축되는 모델에 내장된 편향, 사용된 학습 데이터 세트에 내재된 편향 등
기술적 편향	시스템 훈련이나 의사 결정에 사용된 데이터에 없는 추가적인 편견
사회적 편향	인공지능 시스템의 생애주기 내 활동에서 사회의 역사적 혹은 현재의 불평등을 제대로 고려하지 못하는 것네트워크의 친환경화를 촉진하는 방안 고려

- (프라이버시 및 개인정보 보호) 관련 국내 및 국제 법률, 프레임워크를 통해 프라이버시와 개인정보를 보호해야 함
  - 본 협약은 다섯 가지 주요 프라이버시 보호 요소*를 제시하여 프라이버시가 단순히 정보보호에만 국한되지 않고, 개인의 정체성, 존엄성, 자율성, 심리적, 도덕적 온전성까지 포괄하는 넓은 개념임을 강조함

    * △개인의 사생활과 활동에 대한 접근 제한, △개인사 보호, △개인정보 및 데이터 통제권, △개인적 특징 보호(개성, 정체성, 존엄성, 자율성 등), △개인의 사적 영역과 신체적, 정신적, 도덕적 상태 보호

- (신뢰성) 인공지능 시스템과 생애주기 내 활동에 대한 신뢰성을 증진하기 위해 생애주기 전반에 걸쳐 적절한 품질 및 보안 조치*를 취해야 함

  * 기술 표준, 기술 사양, 보증 기법, 준수 체계 등

  - 기술 표준은 국제 및 국내 인권 기구와의 일관성을 장려하는 투명하고 포괄적인 프로세스를 통해 개발되어야 함

- (안전한 혁신) 인공지능 시스템 생애주기 내 활동의 부정적 영향을 피하면서 혁신을 촉진하기 위해 인공지능 시스템을 개발하고 실험할 수 있는 통제된 환경 혹은 프레임워크를 마련해야 함
  - 인공지능 시스템의 품질, 프라이버시, 인권, 보안, 안전 문제는 개발 초기 단계부터 고려되어야 하며, 일부 위험 요소는 설계 단계에서 효과적으로 관리되어야 하므로 규제 당국의 초기 감독이 필수적임
  - 구체적인 방안으로는 규제 샌드박스와 규제 당국이 새로운 상황에서 인공지능 시스템의 설계, 개발 또는 사용에 대해 어떻게 접근할 것인지 명확히 하기 위한 특별 규제 지침 혹은 무조치 서한이 있으며, 이점은 다음과 같음;

- 인공지능 시스템의 통제된 개발, 테스트, 검증을 허용함으로써 개발 프로세스 초기에 인공지능 시스템과 관련된 잠재적 위험과 문제를 선제적으로 식별 가능함
- 민간 기업, 규제 기관 및 기타 이해관계자 간의 지식 공유가 촉진되어 리소스가 부족한 소규모 기업에 도움이 될 수 있음
- 인공지능 기술은 빠르게 진화하고 있어 기존의 규제 프레임워크가 그 속도를 따라잡기 어려울 수 있으므로 이러한 접근 방식을 통해 인공지능 기술의 발전에 맞춰 규제를 테스트하고 조정하면 더 나은 규제를 수립할 수 있음
- 규제 당국이 인권, 민주주의, 법치주의를 위해 인공지능 기술을 이해하고 감독하는 데 적극적으로 참여하고 있음을 보임으로써 대중과 업계의 신뢰를 높일 수 있음

제4장. 구제책: △구제책, △절차적 보호장치

- (구제책) 인권 보호를 위해 국제 및 국내 법적 체계를 인공지능 시스템에 적용하고, 이러한 체계가 투명하고 효과적인 해결 방안을 제공하도록 보장해야 함
  - 인공지능 시스템의 생애주기 내 활동으로 인한 인권침해가 발생할 경우, 침해 당사자들은 당국에 불만을 제기할 수 있는 효과적인 수단*을 마련해야 함

    * (예) 공익 단체가 불만을 제기하는 방식 등

- (절차적 보호장치) 인공지능 시스템이 인간의 권리에 미치는 영향을 관리하는 절차적 안전장치를 구축해야 함
  - (절차적 보장) 국제 및 국내 인권법의 절차적 보장, 안전장치, 권리가 인공지능 시스템과 관련된 상황에서도 여전히 유효하며, 인권에 영향을 미치는 결정을 내리거나 결정을 지원하는 경우 인간의 감독을 반드시 포함해야 함
  - (인공지능과의 상호작용에 대한 알림 의무) 인공지능 시스템과 상호작용하는 사람들에게 조작이나 기만의 위험을 방지하기 위하여 인공지능과 상호작용하고 있다는 사실이 고지될 수 있도록 해야 하며*, 특정 상황에서는 예외가 있을 수 있음**

    * (예) 정부 웹사이트에서 AI 챗봇과 상호작용하는 경우

    ** (예) 법 집행 시나리오처럼 인공지능 시스템 사용 사실을 공개하면 역효과가 발생할 경우, 인공지능 사용이 명확한 상황

제5장. 위험과 악영향 평가 및 완화: △위험 및 영향 관리 프레임워크

- (위험 및 영향 관리 프레임워크) 인공지능 시스템의 위험과 영향을 평가·관리하는 체계를 구축하여 인권, 민주주의, 법치주의에 미치는 부정적 영향을 방지해야 함
  - 상황에 따라 유연한 규제 방식을 채택할 수 있으며, 인공지능 시스템이 수용 불가능한 위험을 초래한다고 판단되는 경우 모라토리엄이나 금지와 같은 조치를 고려해야 함

제6장. 협약의 이행: △차별 금지, △장애인 및 아동의 권리, △공공 상담, △디지털 리터러시와 역량, △기존 인권의 보호, △더 넓은 보호

- (장애인 및 아동의 권리) 국내법 및 국제 의무에 따라 장애인과 아동 권리 존중의 필요성과 취약성을 충분히 고려해야 함
  - 디지털 리터러시 교육을 통해 아동과 장애인의 권리를 보호해야 하며, 인공지능 기술이 아동 성착취 및 성적 학대에 악용될 위험성 예방에 초점을 맞춰야 함
- (디지털 리터러시와 역량) 디지털 리터러시 교육은 인공지능 기술의 사회적 악영향을 완화할 수 있으므로, 취약 계층과 법원, 국가 감독 기관, 데이터 보호 기관, 인권 기구, 소비자 보호 기관 등 관련 책임자들을 대상으로 실시해야 함
  - 협약의 해설 문서에서 제시한 디지털 리터러시 교육 주제는 다음과 같음;

a 인공지능의 개념
b 특정 인공지능 시스템의 목적
c 다양한 인공지능 모델의 기능과 한계
d 인공지능 시스템의 설계, 개발 및 사용과 관련된 사회문화적 요인(인공지능 시스템 학습에 사용되는 데이터 관련 내용 포함)
e 인공지능 시스템 사용과 관련된 인적 요소
f 인공지능 시스템이 사용되는 맥락과 관련된 도메인 전문지식
g 법률 및 정책 고려사항
h 인공지능 시스템의 부정적인 영향을 불균형적으로 경험한 개인이나 커뮤니티의 관점

- (더 넓은 보호) 기존의 국내법과 국제협약이 인권, 민주주의, 법치주의를 수호하는 데 더 강한 보호를 제공할 경우 해당 보호 조치를 인정하고 보장함

제7장. 후속 메커니즘 및 협력: △당사국 회의, △보고 의무, △국제 협력, △효과적인 감독 메커니즘

- (국제협력) 당사국은 비당사국의 협약 가입을 권장하고, 인공지능이 인권, 민주주의, 법치주의에 미칠 수 있는 영향을 연구하고 정보를 상호 교환해야 하며, 이 과정에서 비당사국과 학계, 산업계, 시민 단체 등의 비국가 행위자의 의견을 반영해야 함
  - 공유하는 정보에는 인공지능이 인권, 민주주의, 법치주의에 미치는 영향을 예방하거나 완화하기 위한 연구와 민간 부문에서 발생한 위험 및 영향에 대한 정보가 포함됨
  - 정보 공유 과정을 통해 인공지능의 다양한 영향을 포괄적으로 이해하고, 부정적 영향을 예방하고 완화하는 것을 목표로 함

제8장. 최종 조항: △협약의 효력, △수정, △분쟁 해결, △서명 및 발효, △가입, △영토적 적용, △연방 조항, △유보, △협약 철회, △통지

- (가입) 본 협약이 발효된 후, 각료위원회의 만장일치 동의를 얻으면 협약 수립에 참여하지 않은 유럽 평의회 비회원국도 협약에 가입할 수 있음

03 소결

- 인공지능 기술로 인한 인권, 민주주의, 법치주의 훼손 사례가 우리나라에서도 발생하고 있음
  - 딥페이크로 인한 인권침해 사례가 다수 보고되고 있으며, 피해자의 상당수가 한국 연예인이고, 일반인 피해 사례 중 미성년자의 비율이 높아 심각성이 큼[8]
  - 제22대 국회의원 선거를 앞둔 1월 말부터 2월 초까지 총선 관련 딥페이크물 약 129건이 삭제되는 등, 우리나라 민주주의에도 부정적 영향을 미치고 있음[9]

8 조수연. "WSJ "한국, 음란물 취약국...딥페이크 피해자 절반 이상 한국인" MBN, 2024.08.30, https://www.mbn.co.kr/news/world/5053471.; 이유나, "'딥페이크' 피해자 10명 중 3명은 미성년자...2년 만에 4.5배 늘어나", YTN, 2024.08.28.,https://www.ytn.co.kr/_ln/0103_202408281009182423

9 정세희, "이재명에 버럭? 尹도 가짜에 당했다…총선 앞 딥페이크 비상", 중앙일보, 2024.03.08., https://www.joongang.co.kr/article/25233767

- 우리나라도 인공지능 기술로 인한 인권, 민주주의, 법치주의 약화를 방지하거나 완화하기 위한 관련 가이드라인과 계획을 마련하고 있으나 법적 구속력은 없음
  - ´20년 12월, 과학기술정보통신부와 정보통신정책연구원은 인공지능 기술로 인한 윤리훼손 이슈에 대응하고 올바른 인공지능의 개발과 활용을 장려하기 위해 '인공지능 윤리기준'을 수립함[10]
  - ´22년 5월, 국가인권위원회는 '인공지능 개발과 활용에 관한 인권 가이드라인'을 통해 인공지능 개발과 활용 과정에서 발생할 수 있는 차별과 인권침해를 방지할 수 있는 가이드라인을 제공함[11]
  - ´23년 2월, 과학기술정보통신부와 정보통신정책연구원은 인공지능의 윤리적 영향을 사전에 평가하여 긍정적 영향은 극대화하고 부정적 영향은 최소화하기 위해 'AI 윤리 영향평가 프레임워크'를 개발하였고, ´24년 시범 시행 중이나, 이를 뒷받침할 명확한 법적 근거는 마련되지 않음[12]
  - ´24년 5월, 과학기술정보통신부는 '새로운 디지털 질서 정립 추진계획'을 발표하였으며, '인공지능 기술의 안전성, 신뢰, 윤리 확보', '딥페이크를 활용한 가짜뉴스 대응', 'AI 생성물 워터마크 의무화' 등의 인공지능 관련 과제가 포함됨[13]
- 인공지능 기술이 인권, 민주주의에 끼칠 피해에 대응하기 위한 몇몇 법안은 이미 통과되었고, 일부는 발의되었지만, 이를 포괄하는 인공지능 기본법은 아직 계류 중이고 인공지능법 제정 및 윤리기준 마련에 대한 국민들의 요구 또한 높아지고 있으므로 신속한 제정이 필요함[14]
  - (통과) 딥페이크 방지법(성폭력범죄의 처벌 등에 관한 특례법 제14조의2)[15]*, 딥페이크 선거운동 금지법[16](공직선거법 제82조의8)**

    * 딥페이크 영상물을 제작·반포 한 자를 '5년 이하의 징역 또는 5천만원 이하의 벌금'으로, 영리를 목적으로 제작·반포 등을 한 경우에는 '7년 이하의 징역'으로 가중처벌

    ** 선거일 전 90일부터 선거일까지 선거운동을 위하여 딥페이크 영상을 제작, 편집, 유포, 상영, 게시하는 행위 금지
  - (발의) 인공지능 발전 진흥과 사회적 책임에 관한 법률안, 인공지능 책임법안, 인공지능 기본법안, 인공지능 개발 및 이용 등에 관한 법률안 등

10 과학기술정보통신부·정보통신정책연구원, "인공지능 윤리 소통채널", https://ai.kisdi.re.kr

11 국가인권위원회. (2024). 인공지능 개발과 활용에 관한 인권 가이드라인

12 과학기술정보통신부·정보통신정책연구원, "인공지능 윤리 소통채널", https://ai.kisdi.re.kr; 양기문(2024.7.31.), "미국과 유럽연합의 AI 영향평가 정책 현황", KISDI Perspectives 2024 July No.3, 정보통신정책연구원.

13 대한민국정부, (2024). 새로운 디지털 질서 정립 추진계획

14 과학기술정보통신부, "AI기술이 인류 위협?…국민 10명 중 6명 "잠재적 이점 더 많아", 대한민국 정책브리핑, 2024.08.07., https://www.korea.kr/news/policyNewsView.do?newsId=148932362

15 성폭력범죄의 처벌 등에 관한 특례법 제14조의2

16 공직선거법 제82조의8

- 본 협약 가입을 통해 우리나라는 인공지능이 인권, 민주주의, 법치주의에 미치는 영향을 국제적으로 협력하여 대응할 수 있음
  - 본 협약은 인공지능의 인권, 민주주의, 법치주의에 미치는 악영향에 대응하기 위한 세계 최초의 법적 구속력이 있는 협약으로, 유럽 평의회 비회원국과 비국가 행위자의 가입도 권장하고 있어 인공지능 분야에서 인권, 민주주의, 법치주의를 수호하는 전세계 표준을 주도할 가능성이 있음
  - 우리나라는 AI 서울 정상회의(´24년 5월)를 개최할 정도로 인공지능 분야에 대한 높은 관심과 적극성을 보이고 있으며, 국제협력의 중요성을 고려하였을 때, 본 협약 가입은 우리나라가 AI 분야에서 글로벌 주도국으로서의 발전에 긍정적인 영향을 미칠 것으로 판단됨
- 또한, 본 협약 가입은 인권, 민주주의, 법치주의를 포괄하는 우리나라의 인공지능 기본법 제정의 기반을 강화하는 계기가 될 수 있음
  - 우리나라는 '서울 선언'(AI 서울 정상회의에서 채택)에서 AI 분야에서 민주주의적 가치, 법치주의 및 인권·기본권 자유와 프라이버시를 보호하고 증진하겠다는 의지를 밝힘[17]
  - 이에 앞서 국가인권위원회는 ´23년 당시, 국회 과학기술정보방송통신위원회에서 논의중인 「인공지능 법률안」과 관련해, 인공지능 개발과 활용 과정에서 인권 침해와 차별을 방지하고 이를 규제할 수 있는 조항 도입의 필요성을 제기함[18]

17 외교부, "AI 서울 정상회의 서울선언 및 의향서", 2024.05.22., https://www.mofa.go.kr/www/brd/m_26779/view.do?seq=537

18 국가인권위원회, "「인공지능 법률안」, '우선허용·사후규제' 원칙 삭제하고, 인권영향평가 도입해야", 2023.08.24., https://www.humanrights.go.kr/base/board/read?boardManagementNo=24&boardNo=7609439&men

참고문헌

공직선거법 제82조의8

국가인권위원회. (2024). 인공지능 개발과 활용에 관한 인권 가이드라인

국가인권위원회, "「인공지능 법률안」, '우선허용·사후규제' 원칙 삭제하고, 인권영향평가 도입해야", 2023.08.24., https://www.humanrights.go.kr/base/board/read?boardManagementNo=24&boardNo=7609439&men

과학기술정보통신부, "AI기술이 인류 위협?…국민 10명 중 6명 "잠재적 이점 더 많아", 대한민국 정책브리핑, 2024.08.07., https://www.korea.kr/news/policyNewsView.do?newsId=148932362

과학기술정보통신부·정보통신정책연구원, "인공지능 윤리 소통채널", https://ai.kisdi.re.kr

과학기술정보통신부·정보통신정책연구원. 2024 인공지능 윤리영향평가 프레임워크

대한민국정부, (2024). 새로운 디지털 질서 정립 추진계획

성폭력범죄의 처벌 등에 관한 특례법 제14조의2

양기문(2024.7.31.), "미국과 유럽연합의 AI 영향평가 정책 현황", KISDI Perspectives 2024 July No.3, 정보통신정책연구원.

이유나, "'딥페이크' 피해자 10명 중 3명은 미성년자…2년 만에 4.5배 늘어나", YTN, 2024.08.28., https://www.ytn.co.kr/_ln/0103_202408281009182423

외교부, "AI 서울 정상회의 서울선언 및 의향서", 2024.05.22., https://www.mofa.go.kr/www/brd/m_26779/view.do?seq=537

정세희, "이재명에 버럭? 尹도 가짜에 당했다…총선 앞 딥페이크 비상", 중앙일보, 2024.03.08., https://www.joongang.co.kr/article/25233767

조수연, "WSJ "한국, 음란물 취약국…딥페이크 피해자 절반 이상 한국인", MBN, 2024.08.30, https://www.mbn.co.kr/news/world/5053471.

Council of Europe, Venice Commission (2011). "Report on the Rule of Law (CDL-AD(2011)003rev).", https://www.venice.coe.int/webforms/documents/?pdf=CDL-AD(2011)003rev-e.

Council of Europe. "What Are Human Rights?", https://www.coe.int/en/web/portal/what-are-human-rights.

Council of Europe . "Democracy.", https://www.coe.int/en/web/compass/democracy#What%20is%20Democracy.

Council of Europe. "Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law." Council of Europe Treaty Series 225, Council of Europe, 2024.

Council of Europe. “Explanatory Report to the Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law.” Council of Europe Treaty Series 225, Council of Europe, 2024.

Council of Europe, “Council of Europe Framework Convention on Artificial Intelligence and Human Rights, Democracy and the Rule of Law brochure.”

Council of Europe. “Key Facts.” https://www.coe.int/en/web/portal/the-council-of-europe-key-facts.

dig.watch (2024.05.20). “Council of Europe adopts first-ever international treaty on AI”, https://dig.watch/updates/council-of-europe-adopts-first-ever-international-treaty-on-ai.

dig.watch. “Convention on AI and human rights (CoE process)”, https://dig.watch/processes/convention-on-ai-and-human-rights-council-of-europe-process.

KISDI Perspectives 발간 내역

KISDI PERSPECTIVES는 국내 외 정보통신방송 관련 주요 정책 및 시장 동향을 분석한 리포트입니다.

문의 : 노희윤 전문연구원 (정보통신정책연구원 방송미디어연구본부, hyoooon@kisdi.re.kr, 043-531-4042)