인공지능과 데이터 거버넌스
- 원칙과 규제 (유럽연합)

스프레드아이 - AI for Everyone, 김정원
jungwon@spreadi.org

목차

참조자료

2024년 1월 21일: 최종안 - EU AI Act Texts - Final Draft in 2024

EY (2023) “Political agreement reached on the EU Artificial Intelligence Act”, EYG no. 011630-23Gbl, 10 December 2023

EU Parliament (2023) “Artificial Intelligence Act: deal on comprehensive rules for trustworthy AI” 9 December 2023

EU Commission (2023) “Artificial Intelligence Questions and Answers”, European Commission, 12 December 2023

임형주 (2023) “인공지능 관련 법 제도의 주요 논의 현황” TTA 저널 특집 스페셜 리포트 207호, 2023년 05/06월호

Veale, Michael, and Frederik Zuiderveen Borgesius (2021) “Demystifying the Draft EU Artificial Intelligence Act”, A Journal of International Law and Technology, SocArXiv. July 6. doi:10.9785/cri-2021-220402. Presentation here.

Gross, Killan, DG CONNECT EU Commission (2024) From Brussels to the Bay: EU AI Act

Kenniscentrum Data & Maatschappij Webinar (2024) Exploring 5 Key Topics under EU AI Act

Gabriele Mazzini, DG Connect, EU Commission (2024) The EU AT Act: Lessons for US Policymakers

Taylor Wessing Germany Webinar (2024) Tracking the EU AI Act, an interview with tech journalist Luca Bertuzzi

Merantix AI Campus seminar (2024) The Finalised EU AI Act: Implications for Businesses, Engineers and Entrepreneurs

Jantina Tammes School, House of Connections, Groningen (2024) Navigating the EU AI Act - Round 1: The general regulatory approach of the EU AI Act

EU 인공지능법 최종 문구확정 - 2024년 1월 21일

배경

세계 최초의 인공지능 종합 규제 법령 초안 제안 (2021년 4월) - 인공지능법 초안의 한글해석본

유럽 연합은 유럽의 인공지능 전략 (2018년 4월), 인공지능 합동 계획(2018년 12월), 신뢰할만한 인공지능 윤리 가이드라인 (2019년 4월), 인공지능 백서(2020년 2월)를 발표하고, 이를 기반으로 인공지능법이 제안됨

유럽연합의 인공지능 법안 발표 이전의 인공지능 윤리 관련 정책과 내용은 소프트웨어정책연구소에서 발간한 '유럽(EU)의 인공지능 윤리 정책 현황과 시사점: 원칙에서 실천으로'에 소개. 인공지능법 발표 이전에 소개된 유럽 연합의 인공지능 윤리 가이드라인과 평가는 법적 의무사항이 아닌 권고사항

Timeline

2021년 4월 21일: 유럽 연합 집행위원회 (EU Commission) 인공지능법 초안 제안
2022년 12월 6일: 유럽 연합 회원국 이사회 (EU Council) 초안 동의, 2023년 6월 14일 유럽 연합 의회 (EU Parliament) 투표로 통과
2023년 12월 9일: 유럽연합 의회, 회원국 이사회, 집행위원회가 최종안의 (정치적) 잠정 합의
2024년 1월 21일: 최종안 문구 확정 (여전히 부록의 세부항목 변경여지 남아있음), 2024년 3월 14일 유럽연합 의회 최종안 투표 가결

관련법안

인공지능법에서 다루지 않은 기타 영역의 데이터와 인공지능 유통 및 로봇 관련 EU의 규제 법안 - **인공지능법은 특히 아래 개인정보보호법에서 명시되지 못한 개인의 권리을 보호하기 위한 보안적 법안으로 제정됨 - 인공지능과 데이터를 다루른 주체는 이 두 법안을 모두 준수하여 시스템을 운영할 것**.

디지털서비스법(Digital Services Act: 연구 목적이나 추천 시스템의 데이터 활용 관련) - 한글 해석본
디지털시장법(Digital Markets Act: AI 관련 하드웨어, OS, 소프트웨어 유통 관련) - 한글 해석본
기계 규제법(Machinery Regulation- 보건,의료, 안전 목적으로 쓰이는 기계 관련) - 한글 해석본
데이터거버넌스법(Data Governance Act: 데이터 공유 프레임워크 관련) - 한글 해석본 등에서 현재 초안이 공개 후 논의 중
일반 개인정보 보호법(GDPR:General Data Protection Regulation): 한글 해석본

EU 인공지능법 - 인공지능 정의

OECD가 정의한 인공지능에 준함

AI 시스템은,
1) 다양한 수준의 자율성으로 운영되도록 설계되어, **배치 후 적응성(adaptiveness)과 자율성**의 수준을 다르게 보일 수 있으며,
2) 명시적 또는 암묵적 **목표를 위해**,
3) 어떤 입력 데이터를 받아 **예측, 콘텐츠, 권고(추천) 또는 결정**과 같은 출력을 생성하는 방법을 **추론하는 기계 기반 시스템**이며,
4) 이러한 **시스템의 출력물은 물리적 또는 가상 환경에 영향**을 미칠 수 있다. 다양한 AI 시스템은 현장 적용시 자율성과 적응성 수준이 달라진다.

참조: OECD AI Principles 2023, https://oecd.ai/en/ai-principles

- 초안에서는 아래와 같은 보다 광범위한 범위로 AI를 정의했었다 - 일반적인 데이터 처리 소프트웨어가 포함되었었다 - 이들은 최종문구에서 이상처럼 자율성, 적응성, 추론의 능력을 갖출때만 AI 시스템으로 간주 . 2023년 12월안에서는 OECD의 정의를 따라 좀 더 정의의 범위가 좁혀졌다. 인공지능 시스템으로는 “머신러닝, 전문가 시스템, 논리 시스템, 베이지안 / 통계 접근법 등을 모두 포함하는 광범위한 소프트웨어 개발 프레임워크”

EU 인공지능법 - 적용범위

AI **개발자, 공급자(수입 및 유통업체 포함), 배치자(Depolyer)** 모두에게 해당됨.

AI 시스템이 운영되는 장소와 AI 시스템이 생성하는 결과물이 EU에서 이용되는 경우 EU역외에 위치한 개발자, 공급자 모두 규제 대상.

이는 EU를 시장으로 비즈니스를 하는 글로벌 모든 기업이 규제 대상이 되는 것을 의미. 고위험군의 경우 AI 구매 후 사용시 배치자에게도 의무사항 부과.

또한 EU 인공지능법은 유럽연합의 법안(Law)이고 지침(Directive)이 아니기 때문에 각 회원국의 법령의 상위법으로 적용됨. 각 회원국 개별의 국가 법령으로 EU 인공지능법에 반대되는 내용은 회원국내에서 법적구속력이 없게됨.

법적용의 예외 대상

- EU 법규 권한 범위 외의 목적으로 사용되는 AI 시스템, 예를 들어 군사나 방위 분야로만 사용되는 AI 시스템
- 과학 연구 및 발견의 목적을 위해 특별히 개발되고 사용되는 AI 시스템
- 시장에 출시되거나 사용되기 전의 AI 시스템에 대한 연구, 시험 및 개발 활동 - 샌드박스 등의 조건에 만족하는 테스트만 허용
- 개인용도로, 비전문적 활동의 일부로 개인이 AI 시스템/서비스를 이용할 때
- EU 역외의 공공기관 혹은 국제기구에서 사용하는 AI 시스템 - 이 경우 관련하여 국제 합의가 명문화되어 있어야 함
- 오픈 소프트웨어 AI : 이 중 **범용(General Purpose) AI와 금지되거나 고위험 AI 시스템으로 분류될 경우는 제외**

EU 인공지능법이 적용되는 대상 - 상세설명

- 공급자 (Providers) :
  - EU 시장에 AI 시스템나 서비스, 혹은 범용 AI 모델을 개발하여 자신의 제품이름과 상표로 EU 시장에 출시하는 주체.
  - 공급자 혹은 AI 시스템이 EU의 역외 위치하지만 (EU 역외에 소재지를 갖고있어도) AI 시스템의 결과물이 EU역 내에서 사용될 경우, 이러한 시스템의 공급자도 EU 인공지능법에 적용되는 공급자

- 배치자 (Deployers)
  - EU에 소재지를 두거나 주소를 갖고 AI 시스템을 자신의 관할하에서 배치-사용하는 주체.
  - 배치자 혹은 AI 시스템이 EU의 역외 위치하지만 (EU 역외에 소재지를 갖고있어도) AI 시스템의 결과물이 EU역 내에서 사용될 경우, 이러한 시스템의 배치자도 EU 인공지능법에 적용되는 배치자

- 유통업자 (Distributors)
  - AI 시스템을 EU 시장에 제공하는 주체

- 수입업자 (Importers)
  - EU 외부의 개인 또는 단체의 이름 또는 상표가 부착된 AI 시스템을 EU 시장에 출시하는 주체

- 제품 제조업자 (Product Manufacturers)
  - 자신의 제품이름과 상표로 AI 시스템을 EU에 제공하는 주체

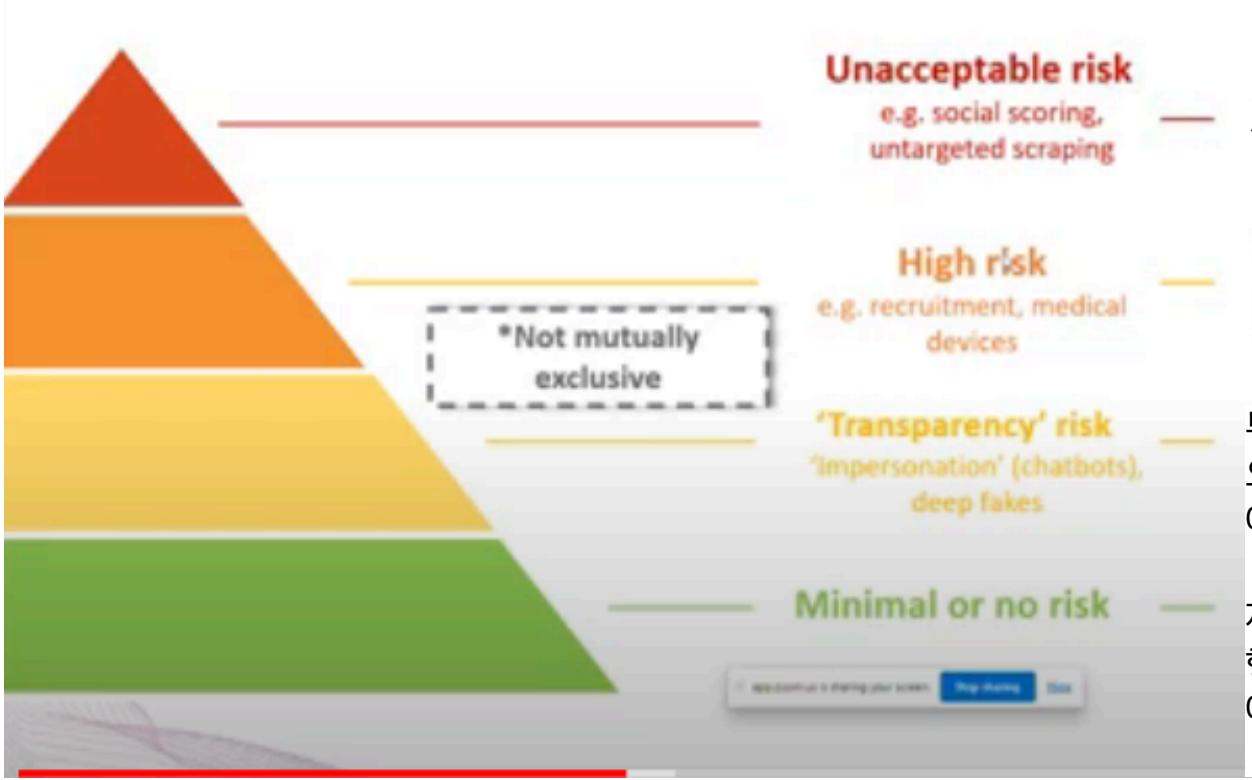
EU 인공지능법 - 접근방법

인공지능 시스템이 제공하는 서비스의 위험 수준에 따라 네 분류(**용인할 수 없는 위험, 고위험, 제한된 위험, 저위험**)로 구분하고, 위험 수준에 비례하는 관리 절차 제안.

AI 시스템의 사용목적, 사용정도, 건강·안전·기본권에 미치는 영향, 피해의 강도 및 범위, 피해복원 가능성 정도 등을 고려하여 AI 시스템의 위험수준을 평가

파운데이션모델, 생성 AI를 포함하는 범용 AI(GPAI:General-purpose AI)는 이러한 위험접근법을 따르지 않음

인공지능 서비스의 위험수준에 따른 규제



사용금지 예) 사회적 평가점수 시스템

적합도 평가 이행의 의무 및
요구사항 준수시 시장출시
예) 사법, 교육/인사 평가 시스템

투명성 준수 및 정보 공유 등의
요구사항 준수시 시장 출시
예)챗봇, 감정인식 시스템

제약없이 시장 출시 -
행동강령 준수를 권고
예) 스팸메일 필터, 사진필터 등

EU 인공지능법 - 용인할 수 없는 금지 AI

기본권 침해 등 EU 가치에 위배되는 아래 목적을 갖는 AI 시스템 활용은 금지 (합법적으로 허가를 받는 연구 목적의 개발 운영만 가능)

- **의사결정을 저해: 잠재의식에 영향을 미치는 기술을 통해 사람들의 행동을 왜곡하고 정보에 입각한 의사결정을 저해하여 정신적 혹은 육체적 위해를 가할 수 있는 시스템**

  예) 트럭운전사의 피곤도가 높아져 운전하지 못할 때 잠을 깨우기 위해 은근히 오디오를 틀어 필요이상으로 위험한 상황에서 운전을 하게 하는 시스템.
  비판) 인공지능의 판단으로 차별적 결과를 가져와 간접적으로 정신적 혹은 육체적 위해를 가하는 시스템의 경우는 포함되지 않음.

- **취약성을 악용: 나이, 신체 또는 정신 장애, 사회적 및 경제적 상황 등과 같은 특정 그룹의 취약성을 악용하여 인간의 자유의지를 회피하여 행동을 조작하는 시스템. 이로인해 정신적 혹은 육체적 위해를 가할 수 있는 시스템**

  예) 보이스 비서 기능이 달린 장난감 인형 - 게임을 가장하여 아동이나 정신적으로 취약한 사람들의 행동을 조금씩 유인하여 위험한 행동을 하게끔 유도하는 시스템
  비판) 분류시스템의 경우 위해를 가하는 경우와 그렇지 않는 경우가 운영에 따라 혼재할 땐? AGI 시스템?

EU 인공지능법 - 용인할 수 없는 금지 AI

- **사회적 점수 시스템**

  사회적 점수 시스템이란 각 개인의 사회적 행동이나 개인의 특성을 기반으로 '신뢰도' 점수를 생성하고,
  '신뢰도' 점수에 기반하여 개인 또는 그룹을 부당하게 혹은 차별하는 대우를 하거나,
  또는 정당하고 비례적이지만 입력 데이터에서 관련 없는 '맥락'에서 해로운 대우를 하는
  최종합의로 추가된 부분: **공공서비스뿐 아니라 민간기업에**서도 유사 시스템을 개발 공급 활용하는 것 금지

  예) 중국의 사회 점수 시스템, 사회아동돌봄이 필요한 아동을 분별하는 시스템에서 부모의 "적절한 돌봄"과 관련이 없는 데이터(예: 이혼, 복지사 면담 결석 등)로 부모를 판단할 때. - 영국 사례 참조

- **사람들의 감정을 추론 인식하는 AI 시스템:**

  특히 직장내 **인사 시스템**이나 **교육기관에서의 입학, 평가등에 사용**되는 시스템

  최종합의안 예외조항: **의료적이거나 안전을 이유로 적용되는 경우는 예외**

EU 인공지능법 - 용인할 수 없는 금지 AI

- **개인 생체인식 데이터를 기반으로** 인종, 성적지향성, 정치적 견해, 노동조합 가입여부, 종교적 또는 철학적 신념 등의 **개인의 민감한 속성을 추론**하여 **사람을 판별하는 시스템** - 다만 합법적으로 취득한 생체인식 데이터의 라벨링이나 필터링, 또는 법집행기관이 생체인식 데이터를 분류하는 경우는 예외

- 인터넷이나 CCTV 영상에서 **무차별적인 안면 이미지 수집**을 통해 안면 인식 데이터베이스를 구축 활용하는 **안면인식 시스템**

- 범죄 예방을 위해 개인의 행동의 범법을 예측하는 **개인범죄행동예측 시스템을 활용한 치안시스템** **최종합의안에서 쟁점이 되었던 부분 -** 개인의 인간적 특성(personal feature)을 기반으로 하는 범죄행동예측 시스템은 금지하지만, 개인의 특성이 아닌 개인과 관련된 다른 배경 특성 등으로 가능한 범법자를 예측하는 시스템은 가능 (예: 전과기록 등)

- **법집행 기관의 '공개적으로 접근가능한 공간'에서 '실시간' 원격 생체인식 시스템 사용.** - 이러한 시스템 활용없이 상당한 피해가 발생할 다음과 같은 범죄 유형에만 예외적으로 사용가능
    - 실종자, 인신매매 피해자, 성착취 피해자를 찾는 경우
    - 생명에 대한 실질적이고 임박한 위협이나 예상되는 테러공격을 방지하려는 경우
    - 중범죄(예: 살인, 강간, 무장강도, 마약 및 불법무기거래, 조작범죄, 환경범죄 등) 용의자 식별의 경우
    - 이와같은 예외 사용경우 다음을 준수해야야 함
        - 경찰은 기본권 영향평가를 완료하고 EU 데이터베이스에 시스템을 등록. 단, 정당한 긴급 상황에서는 등록 없이 배치를 시작할 수 있지만, 부당한 지체 없이 나중에 등록해야 함.
        - 또한 시스템 설치 전에 사법 당국 또는 독립 행정 당국의 승인을 받아야 함. 단 정당한 긴급 상황에서는 승인 없이 설치와 사용을할 수 있지만, 24시간 내에 승인을 요청해야 함. 승인이 거부되면 즉시 사용을 중지하고 모든 데이터, 결과 및 산출물을 삭제해야 함

비판) 왜 '실시간' 원격생체인식 시스템만 금지? 금지 시스템 EU 역외로 수출가능, 법 집행이 아닌 이유의 생체인식 시스템은 이용 가능?

EU 인공지능법 - 고위험 AI

- 고 위험군 AI 시스템은 - 어떤 시스템의 부속으로 존재하는 AI 시스템과 AI 단독 시스템 모두를 포함 -

  1) 인간의 건강에 미치는 영향과 안전성 점검을 받도록 요구되는 시장에 출시되는 상품군에서 AI를 적용하는 경우 - (예: 장난감, 자동차, 의료기기, 기계류, 항공, 자동차, 선박, 철도-기차 등 + 생체인식 시스템, 자연인 분류 시스템 등). 이들 제품을 시장에 출시하기위해서는 **각 회원국에서 지정한 제품 안전성 테스트 전문기관에서 사전 준수성 평가(Conformity Assessment)를 받는 것이 의무화**되어 있음 - Annex II 참조

  2) 이외 인공지능법에서는 **8개의 새로운 고위험군 제품 영역을 지정**하고. 이 경우 안전성 점검 전문 제3기관이 존재하지 않으므로 **제품 개발사가 자체 사전 준수성 평가 수행** - Annex III 참조. 이 경우 회원국간 준수성 평가 표준절차가 일치하지 않는 영역의 경우는 EU 승인 테스트기관에서 평가 실시.

- 부과된 요구/의무사항은 **준수성 평가를 제품 안전성 테스트 전문기관 혹은 제품 공급사가** 시행하여 준수여부를 확인하고, 공급자가 **준수성 평가 통과 마크(EU Conformity Evaulation: CE)를 부착. 인증마크를 획득한 고위험 AI 시스템만 EU에서 운영 가능.** 국가 관할 기관 요청시 **법적 의무사항 준수 입증 등을 의무**화. CE부착 후 법준수여부가 사실이 아닐 경우 벌금 부과

  이들 8개의 고위험군 제품 영역에 해당되지만 **다음과 같은 사용은 고위험군의 의무사항 준수에서 예외**
  - AI 시스템이 직접적인 안전이나 보안에 영향을 미치지 않는 제한된 절차 작업 수행의 경우
  - 이전에 완료된 인간이 제작한 결과물의 품질을 검토하거나 개선하는 경우
  - 최종 의사결정에 영향을 미치지는 않지만 의사결정 패턴(기존 패턴에서의 편차를 발견하여 불일치를 찾아내는)을 감지하는 경우
  - Annext III에 나열된 사용 사례의 목적과 관련된 평가를 위한 준비 작업을 수행하는 경우 - 규제샌드박스 지역과 범위내에서 실제상황을 테스트하는 경우
  - Covid-19과 같은 전염병 등 공공의 이익을 위한 경
  - 이러한 예외의 경우 AI 시스템 공급업체는 시장 출시 혹은 시스템 제공전에 왜 고위험군에서 예외되는가의 의사결정 근거를 문서화해야함

- 시스템 개발자, 공급자, 유통업체, 이용자 각각의 요구사항과 의무사항을 명시
- AI 시스템이 개인을 프로파일링하는 경우(즉, 개인의 업무 성과, 경제 상황, 건강, 선호도, 관심사, 신뢰성, 행동, 위치 또는 이동 등 다양한 측면을 평가하기 위해 개인정보를 자동 처리하는 경우) 항상 고위험으로 간주

EU 인공지능법 - 고위험 AI 목록

1. 용납할 수 없는 금지 AI군에서 제외된 자연인의 생체인식 및 분류 시스템 : 민감하거나 보호되는 속성 또는 특성을 추론하는 생체인식 분류 시스템 및 감정 인식 시스템 - (리모트 시스템, 실시간 이후 스토리지 데이터 분석검색 포함). 단 개인이 주장하는 대로 본인임을 확인하는 생체인식 검증은 제외

2. 시민의 삶과 건강에 위협을 가할 수 있는 중요 인프라의 관리 및 운영에 활용되는 시스템 (예: 도로, 교통 및 물, 가스, 난방 및 전기 공급의 관리 및 운영에 활용되는 AI 시스템)

3. 교육 및 직업훈련:

    (a) 교육 및 직업 훈련 기관에 자연인의 교육기회 결정하는 시스템

    (b) 교육 및 직업 훈련 기관에서 학생의 학습과정을 안내하는 것을 포함한 학습성과 평가 시스템

    (c) 교육 기관 입학에 필요한 시험에서 응시자의 적절한 교육수준 평가 등의 목적으로 사용되는 AI 시스템

    (d) 시험 중 학생행동 모니터링 및 시험금지행동 감지시스템

4. 고용, 근로자 관리 및 자영업에 대한 접근:

    (a) 자연인의 채용 및 선발시 맞춤형 채용광고, 지원서 분석 및 필터링, 지원자 평가 시스템

    (b) 승진, 업무 관련 계약 관계의 종료, 직원의 성격특성 이나 행동특성에 기반한 업무 배정의 과정, 자연인의 성과 및 행동을 모니터링, 평가하는 AI 시스템

EU 인공지능법 - 고위험 AI 목록

5. 필수 민간 및 공공 서비스: (아래는 사례)
    (a) 공공 서비스와 혜택에 대한 자연인의 적격성을 평가하는 AI 시스템. 평가 결과에 따라 공공 서비스와 혜택의 부여, 축소, 취소 또는 회수를 결정하는 AI 시스템
    (b) 자체 사용 및 금융사기탐지를 제외한, 자연인의 금융 신용도를 평가하거나 금융 신용 점수를 결정하는 AI 시스템
        새로 추가된 조항: 건강 및 생명보험의 개인 위험및 보험료 평가시스템
    (c) 의료 지원, 긴급 구조대 서비스의 출동 우선 순위 결정에 사용되는 AI 시스템

6. 법 집행 및 치안:
    (a) 자연인의 공격 또는 재범 위험, 또는 형사 범죄의 잠재적 피해자에 대한 위험을 평가하기 위해 사용하는 AI 시스템
    (b) 거짓말 탐지기 및 유사한 도구로 사용하거나 자연인의 감정 상태를 감지하기 위해 사용하는 AI 시스템
    (c) 딥페이크 탐지 AI 시스템
    (d) 형사 범죄 수사 또는 기소 과정에서 증거의 신뢰성을 평가하는 AI 시스템
    (e) 자연인의 프로파일링을 기반으로 실제 또는 잠재적 범죄 행위의 발생 또는 재발을 예측하거나, 또는 자연인 또는 그룹의 성격적 특성 및 특징, 또는 과거 범죄 행위를 평가하는 AI 시스템
    (f) 형사 범죄의 탐지, 조사 또는 기소 과정에서 자연인의 프로파일링을 위해 법 집행 기관에서 사용하는 AI 시스템
    (g) 법 집행 당국이 데이터에서 알려지지 않은 패턴을 식별하거나 숨겨진 관계를 발견하기 위해 다양한 데이터 소스 또는 대규모 데이터 세트를 조사하는 방식으로 자연인에 대한 범죄 분석에 사용되는 AI 시스템

EU 인공지능법 - 고위험 AI 목록

7. 이주, 망명 및 국경 통제 관리:

(a) 거짓말 탐지를 포함한 자연인의 감정 상태를 감지하기위한 AI 시스템
(b) 입국을 원하거나 입국한 자연인(불법입국과 건강상태 평가)의 보안과 위험을 평가하는 AI 시스템
(c) 자연인의 여행 및 관련 문서의 진위 확인을 위해 사용하는 AI 시스템
(d) 망명, 비자 및 거주 허가를 신청하는 자연인의 적격성 평가, 관련된 소원처리조사를 위해 사용하는 AI 시스템. 여행 서류 검증을 제외한 개인탐지, 인식 또는 식별 시스템

8. 정의와 민주적 과정의 관리:

(a) 사법 당국의 사실 조사, 법 해석 및 구체적 사실에 대한 법 적용, 또는 대체분쟁해결에 도움을 주는 AI 시스템

(B) 선거 및 국민투표 결과나 투표 행동에 영향을 미치는 것. 단, 정치 캠페인 구성, 최적화 및 구조화에 사용되는 도구와 같이 사람과 직접 상호작용하지 않는 출력물은 제외.

EU 인공지능법 - 고위험 AI 준수사항

① **위험관리 시스템 구축**: 시스템 생애주기 전반에 걸쳐 위험관리 시스템 운영되어야 하고, 운영 기록을 남겨야 함.

② **데이터 거버넌스 수행**: 훈련, 테스트, 확인 데이터가 시스템의 원래 목적을 달성하는데 충분히 완전하고, 오류가 없으며, 관련있는 데이터이어야 함. 인종 등의 민감한 개인 데이터는 고위험 AI 시스템 편향성 테스트를 위해서만 공급자(데이터 브로커 제외)만 활용 가능.

③ **기술 문서화:**시스템 아키텍쳐, 알고리즘 디자인, 모델 명세서 등을 포함하여 시스템 공급업체가 제작 보관하고 있어야 함.

④ **자동 로깅 기록**: 시스템 운영의 자동 로깅 기록이 업계에서 인정하는 표준에 따라 수행, 보관되어야 함. 자동로깅시 보관해야할 데이터 목록 제시.

⑤ **이용자에게 투명성 및 정보 제공**: 이용자가 시스템이 만들어내는 결과를 충분히 이해하고, 이용자를 위협하는 사고 발생시 추적 가능한 관련 정보를 공유해야 함. 공유해야할 정보 유형에 따라 공개 대상을 이용자, 공공 DB에 공개, 공급자 보관으로 지정하고 있음. 이용자는 반드시 공급자가 제공하는 매뉴얼에 따라 시스템을 이용하여 시스템 오류나 편향성을 방지할 의무가 있음.

⑥ **사람에 의한 감독:** 사람에 의한 감독이 효과적으로 수행될 수 있도록 인간-기계 인터페이스 도구를 갖추어야함. 특히 생체인식데이터를 다루는 시스템의 경우 시스템 공급업자는 데이터 사용전 반드시 두 명의 자연인이 생체인식 데이터와 대상자간의 매칭에 오류가 없음을 확인한 후 시스템 사용하도록 함. 이와 함께 이 두 자연인의 신상 정보와 이들의 데이터-시스템 확인이라는 의무사항을 미리 시스템에 기록해두어야 함.

⑦ **정확성/견고성/사이버보안**: 위험을 식별하고, 식별된 위험을 적절하게 완화하는 조치를 취해야 함. 이에 시스템이 본래 의도된 목적을 달성하며 일관되게 실행하는지 확인해보아야 함. 이를 위해 평가 항목과 확률적 임계값을 사전 정의하여 위험성을 평가해 봄.

⑧ **EU가 직접 관리하는 고위험군 AI DB에 등록 :** 시장 출시전에 등록해야함. 고위험이 아니라고 자체 판단하는 제공자는 이를 입증할 평가결과를 문서화하고 EU AI DB에 등록해야하며, 당국 요청시 평가 문서를 제출.

고위험 AI 시스템 시장 출시 전 공급자 - 요약

1. **준수성평가** (제43조): 고위험군 AI 준수사항 8가지를 준수하고 있는지 적합도 평가 실시
2. **위험관리시스템** (제9조): **위험 평가 실시 및 위험관리시스템운영** - 고위험 인공지능 시스템과 관련된 **알려진 위험과 예측 가능한 위험을 식별하고 분석**하며, 이러한 **위험을 평가하고 적절한 위험 관리 조치**를 취하는 것이 포함. **위험 평가시 기본권 영향 평가 실시.**
3. **데이터 및 데이터 거버넌스** (제10조): 시스템 개발에 사용된 교육, 검증 및 테스트 데이터 세트가 적절한 데이터 거버넌스 및 관리 관행에 따라야 함 - 데이터가 관련성 있고, 대표성을 갖고, 오류없이 완전하며, 개인정보는 보호되어야 함
4. **기술 문서 작성** (제11조): 시스템의 기술 문서를 작성해야 하는데, 기술 문서는 해당 시스템이 법률의 요구 사항을 준수하는지를 증명해야 하며, 최신 정보를 유지해야 함.
5. **기록 보존** (제12조): 법이 명시하는 적절한 기간동안 시스템이 자동으로 생성된 로그를 보존해야 함.
6. 투명성 및 배치자에 대한 정보 제공 (제13조): 시스템이 운영되는 방식이 충분히 투명하게 되도록 설계되고 개발되어, 배치자가 시스템의 결과를 해석하고 적절하게 사용할 수 있도록 보장해야 함. 공급자는 사용 설명서를 명확하고 이해하기 쉬운 방식으로 제공해야 함.
7. **인간 감독** (제14조): 시스템이 운영 기간 동안 자연인에 의해 효과적으로 감독될 수 있도록 설계되고 개발되도록 보장해야 함. 이는 인간 감독을 가능하게 하는 인간-기계 인터페이스 도구를 구현하는 것이 포함.
8. **정확성, 견고성 및 사이버 보안** (제15조): 시스템이 적절한 수준의 정확성, 견고성 및 사이버 보안을 달성하도록 설계되고 개발되도록 보장해야 하며, 수명 주기 전반에 걸쳐 일관되게 작동되어야 함

고위험 AI 시스템 시장 출시 후 공급자 의무사항 - 요약

1. **시장 출시 후 모니터링** (제61조): 시장 출시 후 모니터링 시스템을 운영하여, 고위험 인공지능 시스템의 성능에 관한 데이터를 적극적이고 및 체계적으로 수집, 기록 및 분석해야 함. 모니터링 시스템은 공급자가 인공지능법의 요구 사항을 지속적으로 준수하는지 평가하고, 시스템 사용 중에 발생하는 모든 위험 또는 문제를 식별할 수 있도록 해야 함.

2. **중대한 사고 및 기능장애 보고** (제62조): 공급자는 고위험 인공지능 시스템의 기본권 침해로 이어지는 중대한 사고 또는 기능 장애로 인해 인공지능법의 의무 위반이 발생하는경우, 해당 사고 또는 법령 위반이 발생한 회원국의 시장 감시 당국에 보고해야 함. 이 보고는 공급자가 시스템과 사고 또는 기능 장애 사이의 인과 관계를 규명한 후에 즉시 제출되어야 하며, 심각한 사고 또는 기능 장애를 인지한 후 최대 15일 이내에 제출되어 함.

3. **보정 조치** (제63조): 공급자가 시장에 출시하거나 용역으로 제공한 고위험 인공지능 시스템이 법에 부합하지 않는다고 판단하거나 그러한 이유가 있는 경우, 해당 시스템을 법에 부합하도록 즉시 필요한 보정 조치를 취해야 하며, 필요한 경우 시스템을 철회하거나 회수해야 함. 또한 공급자는 관련된 고위험 인공지능 시스템의 유통업체 및 적용 가능한 경우 관련 대리인 및 수입업체를 포함한 가치사슬의 다른 운영자들에게도 이를 즉시 통보해야 함.

4. **당국과의 협력** (제64조): 고위험 인공지능 시스템의 공급자는 국가 관할 기관의 이유 있는 요청에 따라 해당 시스템이 법의 요구 사항과 일치하는지를 증명하기 위해 해당 국가 관할 기관에 필요한 모든 정보와 문서를 제공해야 함. 공급자는 고위험 인공지능 시스템과 관련하여 해당 기관이 취하는 모든 조치에 대해 해당 국가 관할 기관과 협력해야 함.

5. **중요한 수정사항**(예: 본 시스템의 목적 변경 또는 규정 준수에 영향을 미치는 변경)에 대해 새로운 준수성 평가 수행:

    - 이는 변경 사항이 원래 공급자나 제3자에 의해 이루어졌더라도 적용됨
    - 한정적 또는 최소한의 위험으로 간주되는 AI 시스템의 경우, 변경 사항 이후에도 원래의 위험 분류가 여전히 적용되는지 확인하는 것이 중요.

고위험 AI 시스템 - 배치자, 수입업자, 유통업자 의무사항 - 요약

시스템 배치자 의무사항 (제29조)

- 배치자는 시스템과 함께 제공된 사용 설명서에 따라 고위험 인공지능 시스템을 사용해야 함. 공공기관과 공공 서비스를 제공하는 민간 기업은 **배치전 기본권 영향 평가를 수행해야함.**
- 배치자가 입력 데이터를 결정하는 경우, 고위험 인공지능 시스템의 목적을 고려하여 입력 데이터가 관련성이 있는지 확인해야 함.
- 배치자는 사용 설명서를 기반으로 고위험 인공지능 시스템의 작동을 모니터링해야 하며, 사용 설명서에 따른 사용이 인공지능 시스템이 위험을 가질 수 있다고 생각할 경우, 공급업체 또는 유통업체에 통보하고 시스템 사용을 중지해야 함. 이 때, 어떠한 이유로 공급업체에 연락을 취할 수 없는 경우, 배치자는 직접 시장감시 당국에 보고해야함
- 배치자는 자동으로 생성된 로그를 보존해야 함.
- 각 회원국 관할 기관의 이유 있는 요청에 따라 해당 기관에 필요한 모든 정보와 문서를 제공하여 고위험 인공지능 시스템이 법의 요구 사항과 일치하는지를 증명해야 하며, 국가 관할 기관이 고위험 인공지능 시스템과 관련하여 취하는 조치에 대해 해당 기관과 협력해야 함

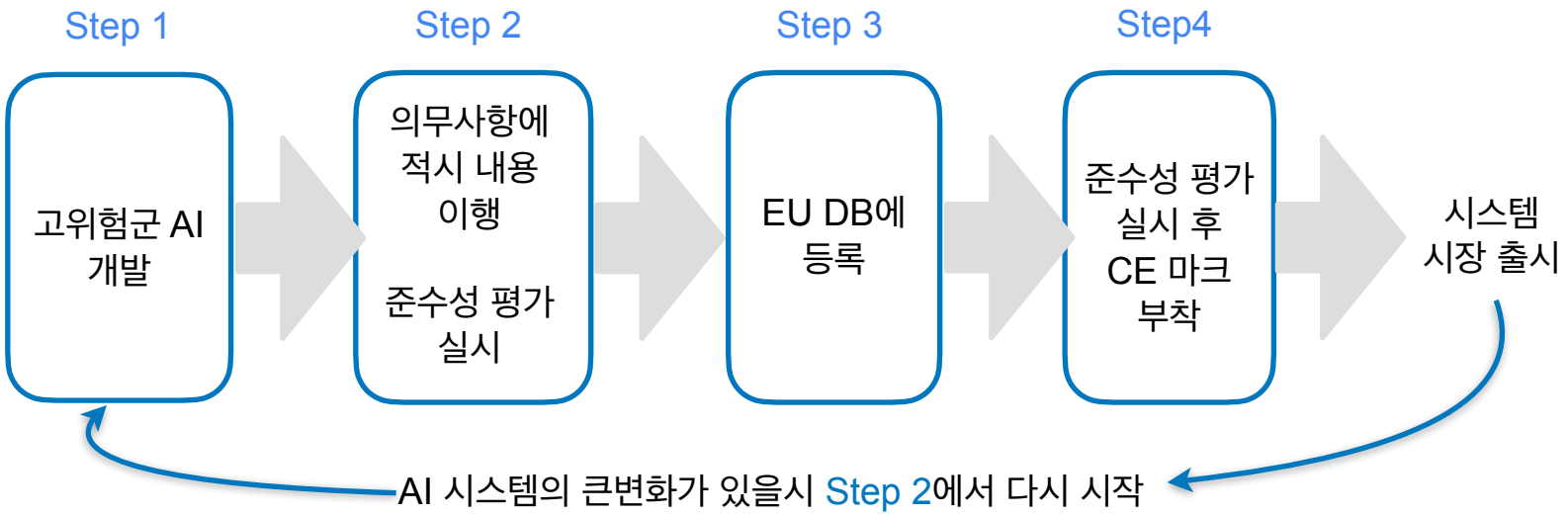
시스템 수입업체 의무사항 (제26조)

- 수입업체는 EU 시장에 제공하는 고위험 인공지능 시스템이 법의 요구 사항을 준수하는지 보장해야 함
- 고위험 인공지능 시스템을 시장에 출시하기 전에 수입업체는 공급업체가 적절한 준수성 평가 절차를 수행했는지 확인해야 함
- 수입업체는 자신의 이름, 등록 상호 또는 등록 상표, 그리고 연락할 수 있는 주소를 고위험 인공지능 시스템에 표시해야 하며, 이를 할 수 없는 경우에는 해당 제품의 포장 또는 동반 문서에 표시해야 함.

시스템 유통업체 의무사항 (제27조)

- 유통업체는 고위험 인공지능 시스템을 시장에 제공하기 전에 시스템이 필요한 CE 일치 표시를 갖추고 있는지, 필요한 문서 및 사용 설명서가 함께 제공되었는지, 그리고 제공업체와 수입업체가 법에서 규정한 의무를 준수했는지 확인해야 함.
- 유통업체가 고위험 인공지능 시스템의 요구 사항과 일치하지 않는다고 판단하거나 그러한 이유가 있는 경우, 해당 시스템을 시장에 제공하기 전에 시스템이 일치하도록 조치를 취할 때까지 제공해서는 안 됨.

EU 인공지능법 - 고위험군의 AI 준수성 평가



AI 시스템이 시장에 출시되면, 시장 감시 기관(Market Surveillance Authority: MSA)이 시스템 제공자, 개발자, 이용자 등의 법이행여부를 모니터링하게됨. 시스템 제공자와 이용자 모두 시스템 출시 후 준수사항을 지켜야함

EU 인공지능법 - 고위험군의 AI 준수성 평가

1. 해당 준수성 평가 절차 결정하기 (부록 VII):
   - 공급자 자체 평가를 하는 고위험 AI 시스템(부록 VII, A장)과 회원국의 관련기관의 준수성 평가가 필요한 시스템인지 확인(부록 VII, B장)
   - AI 법안의 부록 VII를 참조하여 구체적인 시스템 유형에 따라 관련 준수성 평가 절차를 판단.
2. 품질 관리 시스템 수립 및 실행 (제17조):
   - AI 법안의 요구 사항을 준수하는 품질 관리 시스템을 수립, 문서화. 이는 준수성 평가 의무를 포함
   - 품질 관리 시스템에는 고위험 AI 시스템의 설계, 개발, 테스트, 검증, 문서화 및 시장 후 모니터링을 위한 정책, 절차 및 지침이 포함됨
3. 내부 통제 수행 (부록 VII, A장) - 준수성 평가를 자체 실시하는 공급업체는 다음 단계를 수행:
   a) 고위험 AI 시스템의 모든 의무 사항 - 위험 관리, 데이터 거버넌스, 투명성, 인간 감독, 정확성, 견고성 및 사이버 보안 의무 사항을 준수하는지 확인.
   b) 제9조에 따라 적절한 위험 관리 조치를 실시하고 위험 평가를 수행.
   c) AI 시스템의 기술 문서를 작성하여 AI 법의 요구 사항을 준수하는지 증명하여야 함.
   d) 준수성 평가를 실시하여 AI 시스템이 해당 요구 사항을 충족하고 리스크 관리 및 기술 문서가 완전하고 일관되는지 확인
   e) 준수성 평가가 긍정적인 경우, 제48조에 따라 EU 준수 선언서를 작성하고, 제49조에 따라 AI 시스템에 CE 마크를 부착
4. 통보기관 참여 (부록 VII, B장):
   적용 가능한 준수성 평가 절차에 관련기관의 참여가 필요한 경우, 제공업체는 다음을 수행:
   a) 기술 문서 및 기타 필요한 정보를 포함하여 관련기관에 대한 준수성 평가 신청을 제출.
   b) 관련기관과 협력하여 AI 시스템, 시험 시설 및 관련기관이 요청한 추가 정보나 문서에 대한 접근을 제공.
   c) 관련기관이 AI 시스템이 AI 법의 요구 사항을 준수한다고 판단하면, EU 기술 문서 평가서를 발급.
   d) 관련기관의 평가를 고려하여 EU 준수 선언서를 작성하고 AI 시스템에 CE 마크를 부착.
5. 기록 보관 및 당국과의 협력 (제12조, 제61조 및 제64조):
   - 고위험 AI 시스템에서 생성된 로그 및 필요한 보존 기간 동안 기술 문서를 유지.
   - 시장 감시 및 준수성 평가 목적의 관련 기관과 협력하여 필요한 정보 및 접근을 제공.

EU 인공지능법 - 고위험 AI 위험성 평가

1. 알려진 및 예상 가능한 위험 식별 및 분석 (제9조(2)(a)):
   - 시스템의 의도된 목적과 합리적으로 예상 가능한 오용을 고려하여 고위험 AI 시스템과 관련된 위험을 체계적으로 식별하고 분석
   - 이에는 건강, 안전 및 기본권에 대한 위험뿐만 아니라 기술 결함, 부정확성, 편향 및 취약점도 포
2. 위험 평가 (제9조(2)(b)):
   - AI 시스템의 성격, 복잡도, 사용 부문 및 맥락, 피해 가능성 및 불가역성의 잠재적 규모를 고려한 식별된 위험의 확률과 심각성을 추정
   - AI 시스템의 이점 및 대안 솔루션의 가용성과 관련하여 위험 평가
3. 위험 관리 조치 채택 (제9조(2)(c) 및 제9조(3)):
   - 위험 평가를 기반으로 식별된 위험을 최대한 제거하거나 줄일 수 있는 적절한 위험 관리 조치를 채택
   - 이는 설계에 내재된 안전, 위험 완화 및 통제 조치를 우선하고, 잔류 위험에 대한 정보를 제공하는 방식으로 이루어져야 함
   - 위험 관리 조치는 위험의 심각성 및 조치의 기술적 및 경제적 실행 가능성에 비례하여 결정되어야 함.
4. 위험 관리 조치 테스트 및 검증 (제9조(4)):
   - 채택된 위험 관리 조치의 효과를 테스트 및 검증. 테스트 및 검증 결과를 문서화하고 필요시 위험 관리 시스템을 수정하는 데 사용
5. 인간 감독 보장 (제14조):
   - 고위험 AI 시스템 운영 중에 자연인에 의한 모니터링, 탐지 및 개입을 가능하게 하는 인간 감독 조치를 시행.
   - 이는 인간 감독을 위한 도구 및 인터페이스를 제공하는 것뿐만 아니라, 인간 감독을 담당하는 개인들이 필요한 역할을 수행할 수 있는 능력, 교육 및 권한을 갖추도록 보장하는 것을 포함
6. 위험 평가 문서화 및 업데이트 (제9조(4) 및 제11조):
   - 고위험 AI 시스템의 기술 문서의 일부로서 위험 평가 과정과 결과를 문서화. 이는 법령 제11조에 따라 이루어져야 함.
   - 고위험 AI 시스템에 중요한 변경이 발생하거나 새로운 위험이 발생할 때와 같이 필요할 때 위험 관리 시스템을 정기적으로 모니터링하고 검토하며 위험 평가를 업데이트.
7. 품질 관리 시스템 시행 (제17조):
   - AI 법의 위험 관리 요구 사항 및 기타 의무 사항을 준수하는 품질 관리 시스템을 수립, 문서화 및 유지
   - 이에는 위험 평가, 시험, 문서화 및 모니터링을 위한 정책, 절차 및 지침 포함

EU 인공지능법 - 고위험 AI 기본권 영향 평가

1. 관련 기본권 식별 (제9조(2)(a)):
    - 고위험 AI 시스템이 영향을 줄 수 있는 가능성이 있는 기본권과 자유를 식별. 이는 해당 시스템의 의도된 목적, 사용 부문 및 맥락, 영향을 받을 가능성이 있는 개인 또는 그룹의 특성을 고려하여 식별.
    - 관련 기본권은 EU 기본권 헌장에 명시되어 있으며, 인간 존엄성, 개인정보 보호, 차별 금지, 표현의 자유, 어린이의 권리 등의 권리가 포함됨.
2. 기본권 침해 영향 가능성 및 심각성 평가 (제9조(2)(b)):
    - 각 식별된 기본권에 대해 잠재적 부정적 영향의 가능성과 심각성을 평가. 이는 AI 시스템의 성격, 피해의 규모 및 불가역성, 영향을 받을 가능성이 있는 개인 또는 그룹의 취약성 등의 요소를 고려.
    - 평가는 직접적인 영향과 간접적인 영향, 그리고 단기적 및 장기적 효과를 모두 고려해야 함.
    - 평가는 또한 AI 시스템이 기본권을 보호하거나 증진하는 데 기여할 수 있는 가능성을 고려해야 함.
3. 기본권 침해 영향의 필요성과 적절성 평가 (제9조(2)(b)):
    - AI 시스템의 의도된 혜택을 달성하기 위해 기본권에 대한 잠재적 부정적 영향이 필요하고 상당한지를 평가.
    - 더 낮은 위험으로 비슷한 혜택을 달성할 수 있는 덜 침범적인 대안이 있는지를 고려.
4. 완화 조치 식별하기 (제9조(2)(c)):
    - 평가를 기반으로 기본권에 대한 잠재적 부정적 영향을 예방, 완화 또는 해결하기 위한 적절한 조치를 식별합
    - 이는 AI 시스템의 설계나 기능을 수정하거나, 안전장치 또는 통제 조치를 시행하거나, 영향을 받는 개인에게 정보와 투명성을 제공하거나, 보상 및 보상을 위한 메커니즘을 설정하는 것을 포함할 수 있음.
5. 이해관계자와 협의 (제9조(4)):
    - 합당한 경우, FRIA 과정의 일부로서 잠재적으로 영향을 받을 수 있는 개인, 그룹 또는 그들의 대표의 견해를 수렴.
    - 이는 추가적 위험 식별, 제안된 완화 조치의 효과 관련 입력 수집, 영향을 받을 가능성이 가장 높은 사람들의 관점을 고려하는 데 도움이 됨
6. FRIA 문서화 (제11조):
    - 고위험 AI 시스템의 기술 문서의 일부로서 FRIA 과정과 결과를 문서화. 이는 제11조에 따라 시행,
    - 문서화에는 식별된 기본권 위험의 설명, 평가 방법론, 평가 결과 및 제안된 완화 조치가 포함되어야 함.
7. 완화 조치 실시 및 모니터링 (제9조(4)):
    - 고위험 AI 시스템의 전반적인 위험 관리 시스템의 일부로서 식별된 완화 조치를 시행.
    - 완화 조치의 효과를 모니터링하고 AI 시스템의 수명 주기 동안 필요한 경우 FRIA를 업데이트.

EU 인공지능법 - 고위험 AI 준수성 평가 비판

- **현재 많은 전문가들이 가장 강하게 비판하는 부분.** EU에서 산업 표준 프로세스를 개발하는 민간기관인 CEN(European Committee for Standardisation)와 CENELEC (European Committee for Electrotechnical Standardisation)에게 고위험 AI 의무사항 준수 평가 표준 프로세스를 개발하도록 법적 권한을 부여할 예정.

    - 이 경우 고위험 AI 시스템 개발자/공급자는 의무사항 준수 평가를 자체 개발한 절차에 의해 평가하기 보다는 이들 두 기관에서 정의한 표준 평가 프로세스를 따를 가능성이 높아짐 (이들 표준 프로세스를 따르는 것이 의무는 아닐지라도).
    - 이들 표준기관이 정의하는 준수성 평가 절차가 실질적인 법이행의 절차가 될 것으로 보임.
    - 적합도 평가시 정의해야할 법 조항의 모호한 표현들에 대한 평가 통과와 실패의 기준이 표준 프로세스에서 정의될 것. 하지만 이러한 기술 표준이 실제로 법안의 본질적인 목표 - 인간의 기본권 침해 방지-를 구현할 것인지가 의문, 특히 산업 부문과 EU 회원국 역량의 차이, AI 무역을 용이하게하려는 국가와 아닌 국가간의 법안 해석의 차이로 표준개발에 난항을 겪을 수도.
    - 이들 기관은 민간기관으로 산업계의 강한 로비를 거부할 동기가 없고, 시민사회가 표준절차 정의를 모니터링할 기회와 절차가 보장되어 있지도 못함.
    - 이들 기관은 민간기관이므로 이들이 제안할 표준 프로세스에 관하여 유럽의회는 비토권이 없음.
    - 개발사/공급사의 준수성 자체 평가 통과 후 인증부여는 각 회원국의 표준인증 발급 기관에서 이행할 계획이며, 이들 역시 각 국의 공공기관이 표준인증 발급 업무를 아웃소싱하는 민간 표준인증서 발급 전문 기관. 이는 의약품이나 식품 안전성 평가를 직접 실시하고 인증하는 FDA와는 차별되는 지점.
        - 안전성이 요구되는 고위험 시스템의 준수성을 평가하는 제3의 기관이 충분한 자원과 능력을 보유할지도 의문, 시장 출시후 모니터링 집행에 대한 규정이 너무 약하고 사고 발생 후 반응하는 수준임. 좀 더 적극적인 감시와, 시장 감시 당국에게 준수하지 않는 사항을 해결하기위한 보다 강력한 권한을 부여해야함
    - 런던대(UCL) 마이클 빌 교수의 AI 법안 분석 보고서 “Demystifying the Draft EU Artificial Intelligence Act” 참조

EU 인공지능법 - 제한된 위험 AI 의무사항

봇(Bot) 서비스 **공급자 의무사항 (제52조)**

자연인과 상호작용하는 봇과 같은 AI 시스템은 **이와 상호작용하는 자연인에게 인간이 아닌 봇과 상호작용하고 있음을 알려야**함 (이미 이용자가 봇과 상호작용하는 것이 자명하게 이해할 수 있는 상황이나 범죄 예방을 위해 법이 허가한 상황은 제외).이 의무는 봇을 활용하는 플랫폼 배치자가 아닌 봇 시스템 공급자에게 있음.

비판:

공개 API를 활용하여 이용자가 직접 자신의 목적에 맞게 개발 가능한 일반 인공지능(GAI, 예: GPT-3) 시스템의 경우, 개발자, 공급자와 이용자간의 구별을 따로 하기 어려움.

컨텐츠가 무엇인가 또한 어떤 맥락에서 쓰이는 컨텐츠인지에 따라 개발자에게 너무 과도한 부담을 줄 수 있음. 따라서 이런 조항이 실제 이용자 개인 피해를 막는 적절한 보호 의무조항인지, 또한 이러한 의무를 이행하지 않을 때 어떻게 기술적으로 적발이 가능한지도 의문.

범용 인공지능(General-Purpose AI) 규제

2021년 4월에 발표된었던 법령 초안에는 없었던 내용으로 최종안에 추가됨 (제52조)

“GPAI 모델’은 **대량의 데이터로 학습**되고, 그 규모에 맞도록 모델의 **자체 감독을 수행한** 경우를 포함하여, 상당한 **일반성을** 나타내며, 모델이 시장에 출시된 방식과 관계없이 **다양한 고유 작업을 능숙하게 수행**할 수 있으며, **다양한 하위 시스템이나 응용 프로그램에 통합될 수 있는** AI 모델을 의미함. 이는 시장에 출시되기전 연구, 개발 및 프로토타입 활동을 위해 사용되는 AI 모델은 포함하지 않음” - 전형적으로 **매개변수를 약 10억개 이상** 갖는 시스템

적용대상: 파운데이션 모델 개발자(공급자)와 이를 fine-tuning을 거쳐 응용서비스를 제공하는자를 모두 공급자로 간주 (제52c조)

범용 인공지능 모델은 다음의 두 계층의 모델로 구분 - **어떤 층위에 속하는 인공지능 시스템인가에 따라 의무사항이 달라짐**

1) **기본 층위** : 이상의 범용 인공지능의 정의를 따르는 모델

2) **시스템적 위험 층위**: 인공지능 모델 학습시에 사용된 **컴퓨팅 파워의 누적된 양(10^{25} 플로팅 포인트 연산 이상의 컴퓨팅 - ChatGPT, Gemini 등이 해당**)에 따라, 시스템적 위험을 만들어낼 수 있는 고영향 GPAI 시스템으로 구분. **시스템적 위험 층위를 결정 기준은 EU 인공지능 사무소(AI Office)의 모니터링에 따라 지속적으로 변화**할 예정

- 제공업체는 그 모델이 이상의 기준을 충족하는 경우 2주 이내에 EU 위원회에 알려야 함.
  - 제공업체는 이상의 기준을 충족하더라도 모델이 체계적인 위험을 내포하지 않는다는 주장을 제시할 수 있음. 이 경우 EU 위원회는 고유의 판단으로 또는 독립 전문가 과학 패널로부터의 자격이 있는 경고를 받아들여, 모델이 시스템적 위험을 만들어낸다고 결정할 수 있음.

범용 인공지능(General-Purpose AI) 규제

1) **기본 층위 GPAI의 의무사항 (제52c조)**

- 최신 기술 문서 유지 및 관리 - 파운데이션 모델에 기반한 통합서비스까지 투명성 확보를 위한 정보 제공
    - GPAI 모델을 자사의 AI 시스템에 통합할 계획이 있는 **제3의 개발 및 공급업체에게 시스템의 역량과 한계를 포함한 기술 정보 제공**
        - 이 문서는 요청 시 EU AI 사무소 및 국가 당국에 제공되어야 함
    - 충분히 상세한 훈련 데이터 설명과 요약, 테스트 과정과 평가 결과, 저작권보유자가 선택적으로 파운데이션모델에서 (자신의 데이터를) 제외될 수 있도록 하는 메카니즘, 에너지 소비량도 기술 - 이 중 훈련데이터의 충분히 상세한 설명과 요약은 EU AI 사무소의 템플릿을 사용하여 대중 공개하도록함
- EU 저작권법 준수하도록 하는 자체 정책이 갖추어져 있어야 함 - 인터넷에 공유되는 텍스트도 텍스트 권리자가 저작권을 명시해 놓은 경우 AI 시스템 개발자는 텍스트 권리자의 승인없이 데이터 사용금지.
- 오픈 소스모델은 저작권법 준수의 의무와 훈련데이터의 설명과 요약을 공개하는 의무만을 갖는다 - 오픈소스라도 아래 시스템적 위험을 갖는 모델 아래 2) 시스템적 위험 층위 의무사항을 준수해야함
    - 이들 모델은 이미 가중치를 포함한 매개변수, 모델 구조와 사용법이 공개되어 있으며, 또한 모델의 접근, 사용, 수정 및 배포가 공개적으로 허용된 모델

2) **시스템적 위험 층위 의무사항 (제52d조) - 위 기본층위 GPAI 의무사항 준수에 추가로 부과되는 의무사항**

- **위험관리시스템 운영** - 시스템을 시장 출시전 모델 평가 수행 + 체계적인 위험을 평가하고 완화해야함. 모델의 적대적 훈련 수행 (즉, '레드 팀 테스트')
- 체계적 위험 평가의 개요와 취해진 완화 조치를 공개하는 추가적인 투명성 의무를 준수해야 함
- 발견된 심각한 사건과 그에 취해진 수정 조치를 유럽 집행위원회의 AI Office와 관련 국가 관할기관에 문서로 보고
- 적절한 사이버 보안 및 물리적 보호 수준을 확보

모든 GPAI 모델 제공업체는 유럽 표준이 발표될 때까지 자율적으로 실천 기준을 결정하고 이행함으로써 그 의무를 준수하도록 함.
이 실천 기준을 이행하지 않는 제공업체는 대안적인 적합 수단을 위원회 승인을 위해 증명해야 함.

범용 인공지능(General-Purpose AI) 투명성 의무사항

적용대상 : **기본 층위 + 시스템적 위험 층위 GPAI 공급자 모두 (제52조)**

1. 사람들과 직접 상호작용하는 AI 시스템은 상호작용하는 대상에게 그들이 AI와 상호작용하고 있다고 알려주어야 함. 이는 명백한 경우를 제외하고 적용. 범죄 탐지 및 수사를 위해 법으로 허가된 AI 시스템에는 이 규정이 적용되지 않음.
2. 오디오, 이미지, 비디오 또는 텍스트와 같은 콘텐츠를 생성하는 AI 시스템 의 제공자는 출력물에 콘텐츠가 인공적으로 생성되었음을 나타내는 방식으로 표시해야 함. 이는 효과적이고 상호운용 가능하며 견고하고 신뢰할 수 있어야 함. 만약 AI가 소량의 편집만 수행하거나 범죄 탐지 및 수사를 위해 허가된 경우에는 예외.
3. 감정 인식 또는 생체 측정 AI의 배치자는 그것에 노출된 사람들에게 이 시스템이 적용되고 있음을 알려주어야 하며, 개인 데이터를 EU 규정에 따라 처리해야 함. 이 규정은 범죄 탐지 및 수사를 위해 법으로 허가된 시스템에는 적용되지 않음.
4. "딥 페이크"를 생성하거나 조작하는 AI의 배치자는 해당 콘텐츠가 인공적으로 생성되었음을 공개해야 함. 예술적, 풍자적 또는 가상적 작업의 경우 조작의 존재만 공개. 대중에게 공익과 관련된 정보를 제공하기 위해 게시된 AI 생성 텍스트는 인간 검토 또는 편집적 통제를 거치지 않았다면 인공적임을 공개해야 함.
5. 1-4 번 항목의 정보는 최초 GPAI의 상호작용 또는 노출 시에 명확하게 제공되어야 하며, 접근성 요구 사항을 준수해야 함.
6. AI 사무실은 이러한 의무를 효과적으로 이행하기 위해 실천 수칙을 촉진할 것. 실천 수칙이 불충분한 경우, 위원회는 이를 승인하거나 구현을 위한 공통 규칙을 채택할 수 있음.

EU 인공지능법 - GPAI 투명성 의무사항 비판

감정인식 시스템 및 생체데이터 분류 시스템 활용시 **이용자 의무사항 (제52조)**

이상 시스템 활용주체는 범죄 예방을 위해 법적으로 허용된 경우를 제외하고는 이 시스템에 노출되어 상호작용하는 자연인에게 AI 시스템 적용되고 있음을 알려야 함

비판: 생체데이터를 활용한 감정인식이라는 로직 자체가 과학적으로 증명되지 못한 방법임을 고려할 때, 이러한 시스템 자체의 운영 자체를 허용하는 것은 안전하지 못함.

딥 페이크(Deep Fake) **이용자 의무사항 (제52조)**

기존의 사람, 사물, 장소, 또는 기타 개체나 사건을 따라하여 자연인에게 진품으로 보일 수 있는 이미지, 오디오, 비디오 컨텐츠를 생성하거나 조작하는 AI 시스템의 이용 주체는 **최종 컨텐츠에서 보여진 인공적 특성을 공개해야 함.**

이런 공개 의무사항의 예외 경우: 범죄 예방을 위해 사용될 때, 예술과 과학 분야에서 표현의 자유를 위해 사용되는 경우, 개인적으로 프로페셔널하지 않은 목적으로 사용될 때로 규정.

비판: 기술적으로 식별이 어려워 단속에 어려움. 시스템 공급자가 아닌 사용주체의 의무사항을 두어 실질적으로 가공된 컨텐츠임을 인식하지 못한 사용주체가 있을 수 있음. 특성 공개 대상이 너무 광범위하여 불필요한 의무사항이 될 수도.

범용 인공지능(General-Purpose AI) 규제

법적 의무사항의 준수를 지원하기 위한 **실천 지침 개발 및 개발 활동을 할 예정.**

- 국제 접근 방식을 고려할 것
- 현재 요구된 의무뿐아니라 특히 기술 문서에 포함할 관련 정보, 시스템적 위험의 유형 및 성격 및 그 원천, 위험 관리의 방법 등에 관련한 내용을 다룰 것.
- 이는 위험이 가치 사슬 전체를 통해 어떻게 발생하고 구체화될 수 있는지와 특정 도전에 대응하는 방법을 고려할 것.
- EU 인공지능 사무소(AI Office)는 GPAI 모델 제공업체, 관련 국가 관할 기관을 실천 기준 작성에 초청할 계획. 이 과정에 시민 사회, 산업, 학계, 하류 제공업체 및 독립 전문가들이 지원할 수 있음.

EU 인공지능법에서 명시된 다양한 평가 - 요약정리

EU 인공지능법에서 제시된 몇 가지 다른 평가는 무엇이고, 누가 이런 의무를 이행해야 하는가?

	법조항	의무주체	목적	내용
준수성 평가	제16조, 부록7	고위험시스템 제공업체	법 준수여부 확인	EU AI 법에서 명시하는 고위험시스템 의무사항 준수 여부 평가 - 제품 안전성, 사이버 보안, 정확성, 견고성 등의 기술적 요구사항 평가 - 제품출시전 실시 공급자는 평가후 CE마크를 부착하여 본시스템이 EU AI법의 모든 의무사항을 지키고 있음을 알림 - 이후 시장조사기관에서 사실여부를 감독.
위험 평가	제9조 제 1, 2, 3, 4항, 부록 3, 4	고위험시스템 제공업체	위험 요인 식별 및 완화방안제시	준수성 평가의 일부로 요구되고 있음. 고위험시스템의 잠재적 위험을 식별하고 평가. 고위험 시스템의 용도, 알고리즘, 데이터, 프로세스 등을 고려한 위험 분석, 위험완화 및 관리방안제시
기본권 영향 평가	제9조 1,2,4항, 제10조 2-f항, 제 29조 a(2)항, 부록 3,4	고위험시스템 제공업체, 고위험시스템 사용업체	인권보호	고위험시스템의 위험평가의 일부에 포함되는 평가 - 즉 고위험시스템의 위험평가시 안전, 보건 관련 위험이외에도 기본인권을 위협하는 위험도 평가해야함. 고위험시스템이 인권에 미치는 부정적 영향 평가, 프라이버시, 차별금지, 의사자유등의 권리 침해 여부 검토, 인권 보호 및 증진을 위한 조치를 취한 후 이를 기술문서에 명시. 특히, 기본권 영향도 확인은 외부의 인권전문가의 자문을 받을 것을 권고/의무
AI 영향 평가	인공지능법전문(Recital 79) - 권고사항	제공업체, 개발업체, 사용업체	사회/경제/환경에 미치는 영향평가	AI 시스템이 사회, 경제, 환경에 미치는 광범위한 영향 평가, 알고리즘 편향성/투명성/설명가능성/지속가능성 등의 다양한 측면을 고려, AI 활용이 가져온 부정적 영향 및 최소화 및 긍정적 영향 극대화 방안 모색. 시스템 설치와 운영전에 실시할 것을 권고

각 회원국에서의 EU 인공지능법 집행 기관 및 절차

EU 회원국 전체를 아우르는 선제법으로 제안

EU 인공지능법은 각 회원국의 국내 관련법의 상위법으로 적용됨.

비판: EU 회원국이 AI 시스템에 대하여 더 강한 의무조항이나 다른 의무조항을 추가할 수 없음을 의미함.

(예: 실제로 현재 프랑스에서 제안된 디지털 공화국법에서 고위험 AI 시스템 보다 더 넓은 범위의 AI 시스템에 정보 투명성 의무조항을 제안하고 있는데 이를 무력화할 수 있음).

각 회원국에서 상세한 법집행 모니터링

각 회원국 내에서는 시장 감시 기관(Market Surveillance Authority: MSA)을 지정, 현장에서 법 집행 모니터링.

MSA는 규제 감시 기능과 권한을 갖는 정부기관으로, 약 1~25명 직원으로 구성될 것으로 제안.

시스템 이용자와 공급자 모두 법이 규정하는 위험한 상황과 의무사항을 이행하지 못한 것을 MSA에 즉각 알릴 의무

고위험 AI 시스템은 MSA가 관리하는 중앙 DB에 그 목적과 사용법을 등록해야하며, 이 DB 내용은 외부공개.

비판: 법안 적용 대상의 규모와 법안의 복잡성을 고려할 때 현실적으로 너무 작은 규모 제안. 또한 다양한 기술적 전문지식이 요구되는 법 이행 감사기능을 수행할 수 있을지도 의문. AI 시스템이 다양한 부서에서 활용되는 것을 고려했을 때, 경찰, 법원, 복지 예산 부서 등을 관리 감독할 권한을 갖을 수 있을지도 의문. 비즈니스 기밀의 공개 DB화라는 반대 예상.

EU 인공지능법 집행 기관 및 절차

공급업체의 법이행과 각 회원국 관할당국의 법이행 모니터링을 용이하게 하기 위해

- **법령** 준수 자체 평가를 허용하며, **의무 사항은 '조화된 표준'**으로 형식화된 유럽 위원회가 승인한 산업 모범 사례를 사용하여 충족 되도록 유도
- 유럽 위원회는 유럽 표준 기구(CEN 및 CENELEC)에게 '조화된 표준'에서 요구되는 새로운 표준 항목 (특히 고위험 AI 시스템 정의와 준수사항에 상응하는 표준항목)을 제공해줄 것을 공식적으로 요청 (위에서 언급한 고위험 AI 시스템의 시장 진입 의무에 관한 섹션 참조)
- **유럽 표준화 기구는 AI 법령 시행에 필요한 시간 내에 표준을 사용할 수 있도록 목표를 설정** (회원국의 합의된 일정에 따라)하고 있지만, 목표달성을 보장할 수는 없음
- 가능한 경우 유럽 표준화 기구는 국제 표준 기구(ISO 및 IEC)에서 작성한 표준을 최소한으로 수정하여 채택하려고 노력할 것

EU 인공지능법 집행 기관 및 절차

EU내 회원국의 법 집행 상황 모니터링 및 협조사항 결정을 위해 다음의 기구를 설치할 계획

EU 인공지능 이사회(EU AI Board): 회원국의 관련 기관 대표, 유럽데이터 보호 감독관 및 EU 집행위원회의 고위급 대표로 구성.

- AI 법령 및 해당 규정의 원할하고 효과적이며 조화된 시행을 지원
- 위험 요소가 있는 AI 시스템 및 새로운 규칙의 효과적이고 균일한 실행과 관련 의견과 권고를 **EU 집행위원회에 제시**
- **EU 인공지능 사무소에 대한 전략적 감독 (GPAI 모델의 실천 규범의 개발 및 법령 시행을 위한 기술 표준화 활동을 지원 포함)**

EU 인공지능 사무소(EU AI Office) : EU 집행위원회내 기능적으로 독립된 기관

- EU내 인공지능 전문지식과 역량을 보유하여 EU내 중앙집중화된 구조에서 EU 인공지능 법령의 회원국내 실행을 지원
- GPAI 모델의 규칙 시행 지원 - 규칙을 자세히 설명하는 실천 규칙을 작성, 시스템적 위험을 가진 모델 분류, 규칙의 효과적인 시행 및 규정 준수를 모니터링
- **GPAI 모델의 규정 준수 여부 감독** - 체계적인 위험을 가진 모델을 분류하고, 모델 평가 수행, 경고에 대한 조사 및 공급 업체에 수정 조치 요청
  - GPAI 모델 평가가 요구되는 경우 : 1) 정보 요청 권한으로 수집된 정보가 충분하지 않은 경우 준수 여부를 평가하기 위해, 2) 과학적으로 독립된 전문가 패널로부터의 인증된 보고서를 통해 체계적 위험을 조사할 것을 요청받는 경우
- 인공지능 정책과 관련한 EU 기관 및 기구, 또한 이들간의 조정 및 협력 추진.
- 과학 커뮤니티와 강력한 연결고리를 제공하고, 독립 전문가 및 전문 기관을 위한 국제 참조 지점 역할을 하며 전 세계의 유사한 기관과 교류 및 협력을 촉진.

독립 전문가 과학위원회 (Science Panel of Independent Experts): EU 인공지능 사무소 활동을 지원하기 위한 독립 전문가로 구성된 과학 패널

- GPAI 모델의 능력을 평가하기 위한 방법론 개발에 기여. 모델 분류에도 참여하며, 가능한 안전과 위험을 모니터링

자문단 (Advisory Forum): 산업계 (빅테크, 스타트업, 중소기업), 학계, 시민사회대표로 구성. 인공지능 이사회에 기술적 전문지식과 의견을 제공

EU 인공지능법 샌드박스

AI 혁신을 진흥하기 위해 각 회원국내 인공지능법 규제 샌드박스 설치 (제53조)

- 제공업체와 배포업체(예: 중소기업)가 시장에 출시되기 전에 규제 감독 하에서 시스템을 실험, 테스트, 훈련 및 검증할 수 있는 통제된 환경
- 인공지능 사무실은 각 회원국의 관련 당국과 협력하여 인공지능 규제 샌드박스의 설립 및 운영, 자격 및 운영 조건의 명시, 감독 및 모니터링 메커니즘, 책임 및 보상 조치 등을 조정할 것

인공지능 규제 샌드박스 목표 및 일반 조건 (제54조, 제63b조)

인공지능 규제 샌드박스는 다음과 같은 목표를 가져야 함.

a)통제된 실험 및 테스트 환경을 확립하여 혁신을 촉진하는 것; (b) 운영자와 관련 규정 당국 간의 협력을 통해 법적 확신을 강화하는 것; (c) 규제 당국이 인공지능 시스템의 기회와 위험, 그리고 그러한 위험을 방지하고 완화하기 위한 조치의 적합성과 효과에 대한 이해를 향상시키는 것; (d) 기존 규제 프레임워크의 잠재적인 공백에 대한 인공지능 사무실 및 관련 당국에게 정보 제공하는 것.

인공지능 규제 샌드박스 실행조건

- 고위험 AI 시스템의 제공자 또는 예정 제공자가 AI 규제 샌드박스 외부 실제 환경에서 시스템을 테스트할 수 있으며, 이 경우 다음 조건이 충족되어야함.
- 테스트 계획이 2개월 전에 해당 회원국의 시장 감시 당국(MSA)에 제출되고 승인되었어야 함 (제출 후 2개월 이내에 이의를 제기하지 않음).
- 시스템 제공자가 EU에 소재해 있어야 함.
- 데이터 보호 규정을 준수해함.
- 시험이 필요한 시간을 초과하지 않고 최대 2년을 초과하면 안됨. 테스트 중에 발생하는 결과는 시장 감시 당국에 보고되어야 함.
- 제공자 또는 예정 제공자는 시험에 참여하거나 영향을 받는 자에게 정보를 제공하고 명시적인 동의를 얻으며 관련 지침을 제공함.

EU 인공지능법 과징금

1) 용인할 수 없는 AI 시스템의 사용금지 위반 시
최대 3,500만(약 436억 원) 유로, 전년 회계연도 기준 전 글로벌 연매출액의 최대 7% 중 더 큰 금액

2) 고위험군의 의무사항 위반시
최대 1,500만(약 218억 원) 유로, 전년 회계연도 기준 전 글로벌 연매출액의 최대 3% 중 더 큰 금액

3) 인증기관, 관할당국에 부정확, 불완전, 오해의 소지가 있는 정보를 제공한 경우
최대 750만(약 109억 원) 유로 또는 전년 회계연도 기준 전 세계 연매출액의 최대 1% 중 더 높은 금액

4) 중소기업은 위반시 부과된 과징금 금액의 1), 2), 3)에서 언급된 두 기준 중 백분율을 적용하여 부과

EU 인공지능 협정 (EU AI Pact)

EU 인공지능 법령 발효 이전 산업계의 자발적인 인공지능 법령 준수를 위한 준비와 이행을 돕기 위한 활동 (자세한 내용 참조: EU AI Pact)

- EU 집행 위원회가 산업계와 함께할 협정 활동을 발표할 예정
- 주요 EU 및 EU 역외 산업계에서 법령 이행의 모범사례와 방안을 교환하기 위해 2024년 상반기부터 활동 시작 - 관심있는 산업계 참가자는 참가신청을 통해 참가 가능 - 2023년 12월말 현재 - 150여 기관이 참가 등록

EU 인공지능 협정 활동을 통해 기대하는 것

- EU 인공지능법 목표에 대한 공동의 이해
- EU 인공지능법의 구체적인 시행을 이해하고 적응하며 준비 (내부 프로세스 구축, 직원 교육 및 AI 시스템 자가평가)
- 신뢰할 수 있는 AI를 증명하기 위해 마련된 보호장치 관련 지식을 공유하고 가시성과 신뢰성을 높임

향후 일정

일정	법제정 개발 과정
2024년 1사분기	현재 잠정안의 최종 문구 결정 - 결정 후 EU의회와 이사회의 승인
2024년 2사분기-3사분기	현재 잠정안 최종안으로 공포 및 발효예정(2024년 6월 예정)
2024년 2사분기~3사분기 최종안 발효 직후	AI 법령 EU 감독기구인 AI 사무소 설립을 위한 작업 시작. 각 EU 회원국은 AI 규제 샌드박스를 제정
2024년부터 법 시행기간동안	EU 집행위원회는 법적 기한에 맞추어 AI 법령 의무를 이행하기 위해 집행위원회와 자발적으로 협력할 조직과 함께 AI 협정(Pact)을 개시할 것
법효력 발동이후 6개월: 2024년 4사분기~ 2025년 1사분기	용납할 수 없는 AI 시스템 금지조항 발효 및 집행 (Title 1: General Provisions, Title II: Prohibited AI practices)
법효력 발동이후 12개월: 2025년 2사분기~3사분기	파운데이션 모델을 포함한 범용목적 GPAI(General Purpose AI) 관련 의무사항 발효 - Title III, Chapter 4: notified authorities and notified bodies), Title VI: EU AI Board/National Regulators, Title VIIIa : GPAI, Title X : Penalties
법효력 발동이후 24개월: 2025년 4사분기~ 2026년 3사분기	독립적으로 운영되는 고위험군 AI 관련 의무사항 발효 + General Application (고위험군 중 36개월이후로 미루어진 특정 고위험군은 아래 참조)
법효력 발동이후 36개월: 2026년 4사분기~2027년 3사분기	AI 법령 모든 의무사항 발효 - Annex II - High Risk AI system used as a safety component of a product covered by EU Harmonisation legislation