



# 혼돈의 AI Open Source License

LG 전자 최혜성



● 들어가기에 앞서 .....



# 인공지능(Artificial Intelligence) 이란?

---

“사람처럼 학습하고 추론할 수 있는 지능을 가진 컴퓨터 시스템을 만드는 기술”

“인공적으로 만들어진 지능을 가지는 실체, 또는 그것을 만들고자 함으로써 지능 자체를 연구하는 분야”



# 인공지능(Artificial Intelligence) 구현 기술

## Machine Learning

- 머신(기계, 컴퓨터) + 러닝(학습) 이란 ? 기계(컴퓨터)가 학습(스스로 데이터의 규칙을 찾아내는)하는 기술
- 데이터베이스, 레코드 파일, 엑셀 파일 등의 정형 데이터를 처리
- IBM 왓슨, 구글 씨치 알고리즘, 아마존 추천 시스템, 메일 스팸 필터 등

## Deep Learning

- 인간 두뇌의 의사 결정 방식을 모방하는 알고리즘인 인공신경망 방식으로 정보를 처리
- 이미지, 영상, 음성, 텍스트(Text), 소리 등의 비정형 데이터
- 알파고, 자율주행 자동차, 자연어 인식 등

# 인공지능(Artificial Intelligence) 문제점

---

개인정보  
오남용

인간  
윤리와 존엄성

일자리  
감소

# 인공지능(Artificial Intelligence) 문제점 예



Deepfake



젠더리파이  
(성별 식별 AI 플랫폼)



자살 제안하는  
정신과 챗봇



얼굴로 범죄자  
예측하는 AI

# 인공지능(Artificial Intelligence)관련 규제 방안

---

AI에 관한 OECD 원칙에 서명  
(40여개 국가는 AI 시스템을 공정하며 신뢰할 수 있는 방식으로 설계하는데 동의함)

EU는 '신뢰할 수 있는 AI를 위한 윤리 지침'을 제시하고 인공지능 법안(Artificial Intelligence Act)을 발표함

미국에서도 2019년부터 AI를 규제하기 위한 법안을 도입함  
(구글, MS등의 윤리적인 AI 실현을 위한 자율적인 AI 개발 원칙 마련)



● 이제 AI Open Source License를 이야기 해볼게요





# AI Open Source License가 필요한 이유?

---

인공지능이 의도하지 않은 방식으로 사용될 수 있으며 이는 도덕적 딜레마로 이어지고 우리 모두에게 영향을 미칠 수 있어, 이를 제한할 수 있는 License가 필요함

# 기존 Open Source License를 AI에 적용하면 안되나요?

---

Open Source License는 Software/Source Code와 다른 artifact로서의 ML(Machine Learning) Model의 기술적 성격과 기능을 고려하지 않으므로, ML Model 을 보다 책임감 있게 사용할 수 있도록 적절하게 조정된 License가 필요함

개방적(open)이며 책임감 있는(responsible) 있는 ML 개발, 사용 및 액세스를 향한 수단이 될 새로운 라이선싱 프레임워크인 Open & Responsible AI Licenses (OpenRAIL) 있음

# OpenRAIL License의 2가지 주요 기능

## Open

- 로열티 없는 액세스와 유연한 다운스트림 사용 및 재배포, 그리고 파생물의 배포 가능

## Responsible

- License가 부여된 AI Artifact에 사용 기반 제한( Use-based restrictions) 사항을 포함함 (잠재적 사용자가 모델을 오용하는 것을 억제하는 역할)

● RAIL License에 대해 본격적으로 살펴볼까요?

# RAIL(Responsible AI License)란?

---

개발자가 무책임하고 유해한 Application 생성을 막을 수 있게 AI 기술 사용 관련 특정 사용 사례를 제한할 수 있는 행동 조항이 포함된 License



# RAIL(Responsible AI License) Sample License

최소한의 변경으로 적용, 수정 또는 있는 그대로 사용할 수 있는 RAIL License의 Sample

## RAIL-A License (END-USER License )

- Application/Service(Software에 대한 API 기반 원격 액세스를 포함하여 Executable Software Code 또는 Application)에만 적용되는 사용 제한이 포함됨
- 기존 EULA에 END-USER License를 추가하거나, 기존 EULA가 없는 경우에는 END-USER License로 사용함
- RAIL-A License Sample - <https://www.licenses.ai/enduser-license>

# RAIL(Responsible AI License) Sample License

---

## Open RAIL-S License (Source Code License )

- Source Code(AI System 관련 Source Code 및 Script)에만 적용되는 사용 제한이 포함됨
- Source Code의 각 Module block에 License 정보 참조를 포함하는 것이 좋음
- Open RAIL-S License Sample - <https://www.licenses.ai/source-code-license>

# RAIL(Responsible AI License) Sample License

## Open RAIL-M License (Model License )

- Model(Model architecture에 해당하는 학습된 가중치(learned weights)와 매개 변수(최적화 상태 포함)로 구성된 Machine-learning 기반의 어셈블리(checkpoint 포함))에만 적용되는 사용 제한이 포함됨
- Open RAIL-M License Sample - Big Science Open RAIL-M License  
(<https://static1.squarespace.com/static/5c2a6d5c45776e85d1482a7e/t/6308bb4bba3a2a045b72a4b0/1661516619868/BigScience+Open+RAIL-M+License.pdf>)



# BigScience Open RAIL-M License

RAIL(Responsible AI License) 중 Model에 적용되는 Sample License

## 주요 내용 및 주의 사항

- Model 또는 Model 파생물의 사본을 수정 또는 수정 없이 임의의 매체로 복제 및 배포할 수 있음
- Model의 파생물이 다른 License 조건에 따라 출시될 수 있지만 항상 본 License와 동일한 사용 기반 제한을 포함해야 함
- Model, Model 파생물 및 Output은 License의 'Attachment A - USE RESTRICTIONS'(사용 목적 제한 등)에 명시된 제한된 용도에는 사용할 수 없음

## 배포 시 의무 사항

- Model 및 Model 파생물 배포 시, 모든 수취인에게 본 License 사본 제공

# BigScience Open RAIL-M License

## Attachment A-USE RESTRICTIONS

- (a) 적용 가능한 국가, 연방, 주, 지역 또는 국제법이나 규정을 위반하는 방식
- (b) 어떠한 방법으로든 미성년자를 착취, 해하거나 착취 또는 해하려는 목적
- (c) 다른 사람을 해할 목적으로 허위 정보 및 / 또는 내용을 생성하거나 유포하는 행위
- (d) 개인을 해치는 데 사용될 수 있는 개인 식별 정보를 생성하거나 유포하는 행위
- (e) 정보 및/또는 콘텐츠(예: 이미지, 코드, 게시물, 기사)를 생성 또는 배포하고, 정보 및/또는 콘텐츠가 기계 생성이라는 것을 명시적이고 지능적으로 부인하지 않고 모든 컨텍스트(예: 트윗 생성 봇)에 배치
- (f) 다른 사람의 명예를 훼손하거나 폄하하거나 괴롭히는 행위
- (g) 본인의 동의 없이 타인을 사칭하거나 사칭(예: 딥페이크)을 시도하는 행위
- (h) 개인의 법적 권리에 부정적인 영향을 미치거나 구속력 있고 집행 가능한 의무를 만들거나 수정하는 완전 자동화된 의사 결정의 경우

# BigScience Open RAIL-M License

## Attachment A-USE RESTRICTIONS

- (i) 온라인 또는 오프라인의 사회적 행동 또는 알려진 또는 예측된 개인 또는 성격 특성에 근거하여 개인 또는 집단을 차별하거나 해를 입히려는 목적으로 사용되거나 그 효과가 있는 경우
- (j) 연령, 사회적, 신체적 또는 정신적 특성에 따라 특정 집단의 취약성을 이용하여 해당 집단과 관련된 개인의 행동을 신체적 또는 심리적 피해를 유발하거나 유발할 가능성이 높은 방식으로 실질적으로 왜곡하는 행위
- (k) 법적으로 보호되는 특성 또는 범주에 따라 개인 또는 집단을 차별하는 용도 또는 그 효과가 있는 용도
- (l) 의학적 조언 및 의학적 결과 해석을 제공
- (m) 개인이 사기/범죄를 저지를 것이라고 예측하는 것과 같이 사법, 법 집행, 이민 또는 망명 절차의 관리에 사용할 목적으로 정보를 생성하거나 배포(예: 텍스트 프로파일링, 주장 간의 인과 관계 도출) 문서에서 무차별적이고 자의적인 사용

○ Open RAIL-M Sample License가 적용된 Project는?

# StarCoder

BigCode에서 배포한 Code용 LLM(Larger Language Model)

## BigCode의 공동작업 결과물(artifact)

- StarCoder : Code를 위한 최첨단 언어 모델(Language model)- BigCode OpenRAIL-M v1 license
- The Stack : 가장 큰 사전 훈련 dataset - Permissive Code로 구성
- SantaCoder : Small StarCoder로 Python, Java, Javascript와 같은 언어에 대해서만 학습된 Model임
  - (The OpenRAIL license for SantaCoder (<https://huggingface.co/spaces/bigcode/license>))

# BigCode OpenRAIL-M v1 license

StarCoder에 적용되는 License

## 주요 내용 및 주의 사항

- Model 및 Model의 수정본(Modification)을 배포할 수 있음
- 일반적으로 모든 Machine-Learning Model에 적용됨 (Source Code 및 Dataset는 별도의 License 임)

## 배포 시 의무 사항

- Model 및 Model의 수정본 배포 시, 모든 수취인에게 본 License 사본 제공
- Model의 수정본을 공유하는 경우, 수정된 Model의 생성관련 변경 사항을 명확히 기술하는 문서 제공

○ 쉬어가는 코너~

Meta's LLaMa 2 License는 Open Source License 인가?



# LLaMa 2

LLaMA(Large Language Model Meta AI)는 Meta가 공개한 대규모 AI 언어 모델  
(<https://github.com/facebookresearch/llama>)

## 주요 내용 및 주의 사항

- Llama Materials의 파생저작물을 사용, 재생산, 배포, 복사 할 수 있고 Llama Materials의 수정본을 만들 수 있음
- Llama Materials 또는 Llama Materials의 output 또는 result를 사용하여 다른 Large Language Model(Llama 2 또는 그 파생 저작물 제외)을 개선할 수 없음
- Licensee 또는 그 계열사가 제공한 제품 또는 서비스의 직전달 월간 활성 사용자가 7억명이 넘는 경우 별도의 License 요청을 하여 권리를 부여 받아야 함

6.No Discrimination Against Fields of Endeavor(OSI의 Open Source Definition)

5. No Discrimination Against Persons or Groups(OSI의 Open Source Definition)



# LLaMa 2 license

---

## 배포 시 의무 사항

- Llama Materials 또는 그 파생 저작물을 배포 또는 제공하는 경우, Agreement 사본을 제공해야 함
- Llama Materials의 Copy 본에 Attribution Notice를 Notice 텍스트 파일 내에 배포해야 함

## < 참고 자료 >

- <https://www.licenses.ai/ai-licenses>
- <https://blog.opensource.org/metasp-llama-2-license-is-not-open-source/>
- [https://huggingface.co/blog/open\\_rail](https://huggingface.co/blog/open_rail)
- <https://cm.asiae.co.kr/article/2020042910524428652>

