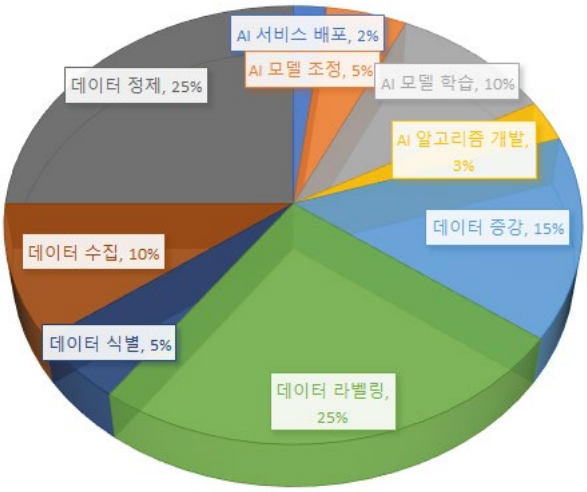
#4. AI 학습 데이터는 특허로 보호받을 수 있을까?

AI 학습 데이터의 중요성

4 차 산업혁명의 핵심기술인 인공지능(AI)은 데이터의 정확도와 양에 따라 그 성능이 좌우됩니다. 우리나라도 글로벌 AI 시장에서 경쟁력을 확보하기 위해 AI 학습에 사용할 수 있는 고품질 데이터 확보를 위한 데이터 댐 구축에 나섰으며, 2020 년부터 2025 년까지 약 2 조 5000 억 원에 달하는 대규모 예산을 투입하고 있습니다. 또한, 최근 통계(아래 그래프 참조)에 따르면, 다수의 기업들도 고품질 데이터를 얻기 위한 '데이터 준비'에 대부분의 시간과 노력을 들이고 있는 것으로 나타났습니다.



[AI 프로젝트에 소요되는 시간 비율]

출처: " [학습용 데이터 구축 ①] 디지털 뉴딜의 핵심 과제, 'AI 학습용 데이터 구축사업'"IT Daily, 김성수 기자, 2021.02.03
"The Ultimate Guide to Data Labeling for Machine Learning", Cloudfactory

특허 대상이 되기 위한 기준

현행 특허법 및 그 심사 기준에 따르면 발명은 물건(장치 또는 시스템) 또는 방법 카테고리로 특정되어야 하며, 컴퓨터 관련 발명의 경우에는 프로그램 기록 매체 또는 데이터 기록 매체의 형태도 허용되고 있습니다. 데이터 구조에 대한 특허등록 사례가 매우 드물게 있지만(아래 참조), 특허 대상 여부에 대한 명확한 판단 기준이 없고 물건 발명인지 방법 발명인지 명확하지 않기 때문에 데이터 구조 및 데이터 세트, 나아가 그 상위 개념인 데이터를 특허 대상으로 청구하는 경우에는 특허법상 발명의 정의 규정에 위배된다고 판단될 가능성이 높을 것으로 사료됩니다.

이벤트 객체에 대응하는 복수의 좌표 정보들을 포함하는 콘텐츠 데이터 구조에서, *상기 이벤트 객체의 기본 정보를 저장하는 제 1 레벨 영역;*

상기 이벤트 객체의 세부 정보를 저장하는 제 2 레벨 영역; 및
상기 좌표 정보들 및 상기 좌표 정보들의 생성 시점 정보들을 대응시켜 저장하는 제 3 레벨 영역을 포함하되,
상기 이벤트 객체의 상기 기본 정보는,
이벤트 종류 정보, 이미지 인덱스 정보, 상기 이벤트 객체와 연동되는 미디어의 재생 시점 정보 및 상기 이벤트 객체의 세부 정보에 대한 이벤트 인덱스 정보를 포함하고,
상기 이벤트 객체의 상기 세부 정보는,
*상기 이벤트 인덱스 정보, 상기 좌표 정보들 중 시작점 좌표 정보, 상기 좌표 정보들 중 종료점 좌표 정보, 상기 이벤트 객체의 영역 정보, 이벤트 색상 정보, 점 굵기 정보 및 상기 이벤트 객체에 대한 수정 발생 유무를 나타내는 수정 정보를 포함하는 것을 특징으로 하는 **콘텐츠 데이터 구조**.*

(한국등록특허 제 10-2003318 호의 청구항 8)

결론

AI 학습을 위한 분석 및 가공된 특정 구조의 데이터는 데이터 요소 사이의 상호 관계를 표시한 논리적 구조로 나타낸 알고리즘으로써, '자연법칙을 이용한 기술적 사상'에 해당되므로 특허법상 보호 대상으로 보는 것이 타당하다고 판단되지만 아직까지 이에 대한 명확한 규정이 없으므로, 데이터를 생성하는 방법, 데이터를 이용한 시스템 등과 같이 다소 우회적인 방식으로 보호하는 것이 안전하며, 데이터 자체를 특허권으로 보호하기는 쉽지 않을 것으로 보입니다.

대안으로, 데이터 부정사용행위를 부정경쟁행위로서 신규 도입한 부정경쟁방지 및 영업비밀보호 관한 법률(이하, "부정경쟁방지법", 2022 년 4 월 20 일 시행)에 기초하여 데이터를 보호하는 방안을 고려해 볼 수 있겠습니다.

부정경쟁방지법상 보호 대상이 되는 데이터는 (1) 특정인 또는 특정 다수에게 제공되는 것이고 (2) 전자적 방법으로 상당량 축적·관리되고 있으며 (3) 비밀로서 관리되고 있지 아니한 기술상 또는 영업상의 정보를 만족하는 것으로서, 아래의 행위에 대해서는 데이터 부정사용행위로서 취급되게 됩니다.

(1) 접근 권한이 없는 자가 절취·기망·부정 접속 또는 그 밖의 부정한 수단으로 데이터를 취득하거나 그 취득한 데이터를 사용·공개하는 행위,
(2) 데이터 보유자와의 계약 관계 등에 따라 데이터에 접근 권한이 있는 자가 부정한 이익을 얻거나 데이터 보유자에게 손해를 입힐 목적으로 그 데이터를 사용·공개하거나 제3자에게 제공하는 행위,
(3) 상기 (1) 또는 (2)가 개입된 사실을 알고 데이터를 취득하거나 그 취득한 데이터를 사용·공개하는 행위, 또는
(4) 정당한 권한 없이 데이터의 보호를 위하여 적용한 기술적 보호 조치를 회피·제거 또는 변경하는 것을 주된 목적으로 하는 기술·서비스·장치 또는 그 장치의 부품을 제공·수입·수출·제조·양도·대여 또는 전송하거나 이를 양도·대여하기 위하여 전시하는 행위

데이터 부정사용행위에 대해서는, 법원에 부정경쟁행위 금지 및 손해배상 등을 청구할 수 있으며, 특허청 등에 행정조사를 신청하여 시정권고 및 공표 등의 구제 조치도 받을 수 있습니다.

본 자료에 게재된 내용 및 의견은 일반적인 정보제공만을 목적으로 발행된 것이며, 특허법인 세움의 공식적인 견해나 어떤 구체적 사안에 대한 의견을 드리는 것이 아님을 알려 드립니다. Copyright ©2022 SEUM IP

윤영진 변리사/외국변호사
Patent Attorney Foreign Attorney/Partner
yjyun@seumip.com