

특허청, 인공지능 학습용 지식재산 데이터 7종 무료 개방

- 4월 29일부터 특허정보활용서비스(KIPRIS^{plus})를 통해 개방 -
- 민간 인공지능 학습 지원 및 지식재산 정보서비스 산업 혁신 기대 -

특허청(청장 김완기)은 4. 29.(화)부터 정보서비스 산업 활성화를 위해 인공지능(AI) 학습에 활용 가능한 지식재산 데이터 7종을 특허정보활용서비스(KIPRIS^{plus}, plus.kipris.or.kr)를 통해 무료로 개방한다고 밝혔다.

특허정보활용서비스(KIPRIS^{plus})는 특허청이 운영하는 대표적인 지식재산 분야 공공데이터 개방 플랫폼이다. 국내는 물론 미국, 유럽, 일본 등 총 13개국에서 발간된 지식재산권 관련 공보와 특허 행정정보 등 총 133종의 지식재산 데이터를 기업과 연구기관, 대학교 등 다양한 수요자들이 자유롭게 활용할 수 있도록 지원하고 있다.

이번에 개방되는 데이터는 특허청의 「2023년 AI기반 특허행정 혁신 사업」 추진 과정에서 구축된 것으로 ▲국내 특허·실용신안 공개공보 ▲미국 특허 공보 ▲의견제출·거절결정서 ▲특허 패밀리 한영 기계번역 ▲국제특허분류(CPC) ▲특허상담 기계독해 ▲특허고객 상담사례집이 포함된다.

최근 많은 기업들이 인공지능 성능 향상을 위해 다양한 데이터를 학습시키고 있으며, 이 과정에서 특허 데이터에 대한 수요 또한 꾸준히 증가하고 있다. 하지만, 특허 분야는 복잡한 구조와 전문용어의 특성으로 인해 인공지능 학습에 적합한 데이터를 분석·구축하는 데 막대한 시간과 비용이 소요되어, 관련 기업들은 양질의 학습데이터를 확보하는 데 어려움을 겪고 있는 실정이다.

이에, 특허청은 민간의 학습데이터 구축 부담을 경감하고, 인공지능을 활용한 정보서비스 개발을 지원하기 위해, 지식재산 데이터를 가공한 고품질 학습데이터를 전면 개방하였다.

이번에 개방하는 학습데이터 7종은 특허청의 행정업무 및 민원 처리 과정에서 생성된 특허정보를 인공지능 기술 학습에 적합하도록 가공한 것으로, 관련 기업과 연구기관은 특허문헌 분석서비스, 인공지능 챗봇 등 다양한 인공지능 기술 개발에 바로 활용할 수 있을 것으로 기대된다.

특허청 최일승 산업재산정보국장 직무대리는 “이번 데이터 개방은 지식재산 정보서비스 산업 고도화에 실질적인 밑거름이 될 것”이라며, “앞으로도 다양한 형태의 특허데이터를 지속적으로 구축·개방하여 지식재산 정보 생태계를 더욱 활성화해 나가겠다”고 밝혔다.

※ 붙임: 특허정보활용서비스(KIPRIS^{Plus}) 개요

담당 부서	산업재산정보국 산업재산데이터관리과	책임자	과 장	이동영 (042-481-5134)
		담당자	사무관	최진아 (042-481-5359)
		담당자	주무관	김성준 (042-481-5137)



더 아픈 환자에게 양보해 주셔서 감사합니다
가벼운 증상은 동네 병·의원으로



□ 서비스 소개

- 특허청이 개방 중인 모든 IP 정보를 실시간으로 활용할 수 있는 ①벌크 데이터, ②Open API 방식의 정보제공 서비스



□ 데이터 제공 형태

- (벌크데이터) 지식재산 정보를 온라인에서 대용량으로 일괄 제공하는 방식
- (Open API) 특허청 DB의 지식재산 정보를 네트워크 연계 기반으로 자체 DB 구축 없이 이용하는 방식

※ '24년 국내 공보, 행정, 부가 정보(10억 건), 해외 공보, 부가 정보(1.6억 건) 제공

□ 데이터 개방 실적(누적)

'23년(119종)	'24년(126종)	'25년(133종)
의약품 실험데이터 미국특허 국문 번역문 베트남 특허공보 인도 특허공보	표준화 특허공보 국내외 권리자변동정보 3종 중국특허 한글번역문 CPC-KPC 분류코드 법적상태정보(특허·실용신안)	인공지능 학습용 데이터 7종

□ 활용 분야

구분	활용 기관	활용 목적
IP관리	특허법인, IP 관리기업, R&D 연구소	IP 고객관리 및 성과 관리에 활용
IP금융	은행, 신용보증기관, 금융 공사	기술 담보의 여신 및 금융심사에 활용
IP평가	기술 평가 전문기관, 평가 연구기관	(유망) 기술 평가 측정에 활용
IP거래	IP 거래 전문기관, 대학	권리 거래를 위한 중개에 활용
IP서비스	IP 번역기업, IP 정보서비스기업	번역 엔진 성능 향상에 활용, 신규 서비스 창출에 활용