

보도시점 2025. 3. 24.(월) 16:00
(2025. 3. 25.(화) 조간)

배포 2025. 3. 24.(월) 07:30

개인정보위, 오픈소스 기반 인공지능(AI) 스타트업 생태계 혁신 지원한다

- 개인정보위, 인공지능 중소·스타트업과 만나 발전 방안 논의

‘딥시크(Deepseek)’의 등장으로 사회적 관심이 높아진 오픈소스 AI 생태계가 국내 AI 스타트업 경쟁력에 미치는 함의와 파급력을 짚어보고, 국내 AI 산업 발전을 위한 정부 차원의 적극적 지원을 약속하는 자리가 마련되었다.

고학수 개인정보보호위원회(이하 ‘개인정보위’) 위원장은 3월 24일(월) 16시, 서울 강남구 소재 스타트업얼라이언스 엔스페이스(&Space)에서 AI 스타트업 관계자들을 만나 현장의 애로·건의 사항을 청취하고, 오픈소스 기반의 국내 AI 생태계 발전 방안을 논의하였다.

오픈소스는 프로그램 개발 시 필요한 소스코드나 설계도를 누구나 확인할 수 있도록 공개하는 방식을 말하며, 이는 비용 부담을 줄이면서도 고성능 AI 모델에 누구나 접근할 수 있도록 하여 과학 기술 발전과 응용 서비스 창출에 큰 기여를 할 수 있다. 특히, 오픈소스는 대규모 AI 인프라는 부족하나 보건의료, 금융 등에 양질의 데이터가 축적돼 있고, 우수한 AI 인력을 보유한 우리나라에 기회를 마련해줄 수 있다. 다만, 추가학습, 검색증강생성*(RAG, Retrieval-Augmented Generation) 등을 거쳐 상용화하는 과정에서 개인정보 처리가 수반될 수 있어 주의를 기울일 필요가 있다.

* 대규모 언어모델(LLM)의 성능 개선을 위해 신뢰할 수 있는 외부 데이터베이스 등을 검색하여, 보다 정확하면서도 최신의 답변을 출력할 수 있도록 하는 기술

개인정보위가 간담회에 앞서 진행한 간이 설문조사에 따르면, 설문에 참여한 10개 기업 중 6개 기업에서 오픈소스 모델에 기반한 응용 서비스를 출시한 바 있고, 오픈소스 모델을 자체 보유한 이용자 데이터 등으로 추가학습하거나, 검색증강생성(RAG)을 통해 보강하여 성능 개선에 활용한다고 답하였다.

이번 간담회에 참석한 국내 AI 스타트업 관계자들은 오픈소스 모델을 적극적으로 연구·활용한 주요 성과와 사례 등을 공유하였다.

먼저, **하주영 스캐터랩 변호사**는 오픈소스 모델과 제반 기술을 적극적으로 연구하여 장점을 흡수하기 위한 노력을 설명하고, 구글 젠마, 딥시크 등 글로벌 오픈소스 모델이 국내 AI 생태계에 미치는 파급력에 대하여 발표하였다.

이어서 **임정환 모레 AI 사업 총괄**이 한국어 답변 성능 강화에 초점을 맞춰 서비스 중인 자사의 언어모델을 소개하면서 서비스 개발·운영 과정에서 느낀 경험을 토대로 오픈소스의 장점과 이로 인해 발생할 수 있는 프라이버시 위협 등에 대하여 발표하였다.

마지막으로 **이재원 엘리스그룹 CISO**(정보보호 최고책임자)는 클라우드 서비스 보안인증(CSAP IaaS)을 획득한 자사 제품을 소개하면서, AI 클라우드 인프라 제공 과정에서 오픈소스 모델을 활용한 사례를 소개하였다.

발표 이후 진행된 자유 토론에서는 간담회에 참석한 기업들이 생성형 AI 개발·도입 과정에서 경험한 데이터 및 개인정보 관련 다양한 애로·건의 사항을 제시하였다.

다수 기업은 자사 또는 고객사가 보유한 이용자 데이터를 AI 개발에 활용하는 경우 발생하는 법적 불확실성 문제에 애로가 있다고 밝혔다. 이를 위해 적절한 이용자 데이터 활용을 위한 명확한 법적 기준 안내, 익명·가명 데이터 처리를 위한 구체적 방법론과 비식별데이터에 대한 재식별 평가기준 마련 등 다양한 건의사항이 제기되었다.

이 자리에서 개인정보위는 AI 신산업 발전을 지원하고 현장 불확실성 해소를 위해, ‘원칙 기반 규율’ 하에서 구체적 데이터 처리 기준*을 제시한 사례 등을 소개하였다. 이어 데이터 활용 장벽 해소를 위해 최근 제3차 국가 인공지능위원회(‘25.2.20.)를 통해 발표한 「AI 데이터 확충 및 개방 확대방안**」의 주요 내용도 설명하였다.

* ▲비정형데이터(‘24.2.), ▲웹 크롤링 데이터(‘24.7.), ▲자율주행기기 촬영정보(‘24.10.) 처리 기준, ▲합성데이터 유용성·안전성 평가기준(‘24.12.), ▲AI 프라이버시 리스크 관리모델(‘24.12.)

** 개인정보위 누리집(<https://www.pipc.go.kr>)에서 발표자료(PPT) 등 범정부 종합 대책 내려받기 가능

개인정보위는 이번 간담회 논의 결과를 바탕으로 국내 AI 확산 추세에 맞추어, 중소·스타트업 등에게 실질적인 도움이 될 수 있는 개인정보 관점에서 맞춤형 ‘생성형 AI 도입·활용 안내서’를 마련할 계획이다.

고학수 개인정보위 위원장은 “우리나라의 경쟁력 있는 AI 혁신 생태계 발전을 위해서는 오픈소스 이점을 최대한 활용할 필요가 있다.”라며, “국내 기관기업에서 오픈소스 AI를 도입·활용하는 과정에서 AI·데이터 처리와 관련된 리스크 요인이 최소화될 수 있도록 중소·스타트업 업계와 긴밀히 협력해 나가겠다.”라고 밝혔다.

담당 부서	개인정보정책국 인공지능프라이버시팀	책임자	팀 장	구민주 (02-2100-3073)
		담당자	사무관	최주은 (02-2100-3072)
			주무관	최홍석 (02-2100-3167)



□ 행사 개요

- (일시/장소) '25. 3. 24.(월) 16:00~17:30 / 스타트업얼라이언스
※ 서울특별시 강남구 봉은사로 215 KST빌딩 3층, 스타트업얼라이언스(&스페이스)
- (주요 참석자) 15명 내외
 - (개인정보위) 위원장, 정책국장, 인공지능프라이버시팀장 등
 - (외부) 인공지능 스타트업 관계자 10명 내외*
- * 뤼튼, 포티투마루, 모레, 인티그리트, 엘리스, 스캐터랩, 로앤컴퍼니, 유니바, 에이슬립, S2W
- (주요 내용) 국내 AI 스타트업의 경쟁력 확보 지원 방안을 모색하고 오픈소스 적용 사례·성과* 공유 및 애로·건의 사항 청취 등

< 주요 발표 내용 >

- ① 딥시크 등 오픈소스 모델의 국내 AI 생태계 파급력(스캐터랩)
- ② 오픈소스 모델을 적용한 개발 사례와 개인정보 아젠다(모레)
- ③ AI 인프라 제공 과정에 오픈소스 모델 적용 사례(엘리스)

□ 세부 계획

※ 사회 : 인공지능프라이버시팀장

시간	주요 논의내용	비고
16:00~16:05	▪ 참석자 소개(5')	사회자
16:05~16:10	▪ 기념 촬영(5')	참석자
16:10~16:13	▪ 인사 말씀(3')	위원장
16:13 ~ 16:37	▪ 발표(24')	
	① (발제1) 딥시크 등 오픈소스 모델의 국내 AI 생태계 파급력(8') ※ 딥시크 오픈소스 워크 기반	스캐터랩
	② (발제2) 오픈소스 모델을 적용한 개발 사례와 개인정보 아젠다(8')	모레
	③ (발제3) AI 인프라 제공 과정에 오픈소스 모델 적용 사례(8')	엘리스
16:37 ~ 17:27	▪ 자유 토론 및 애로·건의사항 청취 등(50')	참석자
17:27 ~ 17:30	▪ 마무리 말씀(3')	위원장

□ 조사 개요

- (대상) 생성형 인공지능 활용 중인 국내 10개 기업이 설문에 참여
- (기간) '25.3.10. ~ '25.3.18. (9일간)

□ 조사 결과

■ (요약) 생성형 인공지능(AI)을 활용 중인 국내 10개 기업이 설문에 참여함.
10개 중 9개 기업은 내부연구·성능개선·응용 서비스 등을 위해 오픈소스 모델을 활용한 경험이 있었고, 6개 기업은 오픈소스 모델을 자체 보유한 이용자 데이터 등으로 추가학습하거나, 검색증강생성(RAG) 등을 통해 보강하여 성능 개선에 활용한 경험이 있었음.

- 오픈소스 활용 경험이 있다고 응답한 9개 기업은 평균 3개(최대 6개, 최소 1개)의 오픈소스 모델을 활용한 경험 보유

※ 본 설문조사 결과는 응답 내용을 바탕으로 요약한 것으로, 실제 운영 상황과 다를 수 있음

- 생성형 AI 모델 도입·이용 형태 : 클라우드 기반 AI 개발 플랫폼 대비 API 제공 모델 및 오픈소스 모델의 활용도 높음
 - ※ ▲ChatGPT(API) > ▲Claude(API) > ▲LLaMA(오픈소스) > ▲Deepseek(오픈소스) > 기타(Mistral, Qwen, Phi, Kanana Nano 등) 순으로 높은 활용도
- 오픈소스 모델 활용 방식 : 오픈소스 모델을 그대로 도입하기보다 자체·고객사에서 보유한 데이터를 통해 추가 학습 후 활용하는 추세
- 생성형 AI 기반 타겟 시장 : B2B 및 B2C 시장이 우세하나, AI의 전방위적 확산으로 공공 분야를 대상으로 한 B2G 시장도 존재
- 국내 AI 도입·활용 전망 : 개념 증명*(PoC) 단계를 지나, 주력 제품·서비스에 적용하고자 하는 수요가 증가할 전망이 우세

* 개념 증명(Proof of Concept) : 특정 기술이나 아이디어가 실제로 구현 가능한지 검증하는 과정