

## 과기정통부, 자율주행 산업 발전 가속화를 위한 국내 최초 'E2E 학습데이터 가이드라인' 발표

2026.06.25

과학기술정보통신부(이하 '과기정통부')는 지난 6월 19일 자율주행 인공지능(AI) 개발에 필수적인 학습데이터를 국내 산·학·연이 함께 구축하고 쉽게 상호 공유·활용할 수 있도록 지원하기 위한 「자율주행 E2E 데이터 구축 가이드라인 및 규격 정의서」(이하 '가이드라인')를 발간하였다고 밝혔습니다.

아래에서는 이번 가이드라인의 추진 배경과 주요 내용, 그리고 그 시사점 등에 대해 살펴보도록 하겠습니다.

### 1. 추진 배경 및 의의

최근 글로벌 자율주행 패러다임은 대량의 데이터를 학습한 하나의 인공지능(AI)이 인지·판단·제어를 통합 수행하는 엔드투엔드(End-to-End, 이하 'E2E') 방식으로 전환되고 있습니다. E2E 방식은 센서 입력부터 제어 명령까지를 하나의 AI 모델로 구현하는 것으로, 정의한 규칙에 따라 움직이는 Rule-based 방식과 달리 사전에 정의되지 않은 상황에서도 유연하게 대응할 수 있다는 특징이 있습니다. 다만 이러한 방식은 막대한 양의 학습데이터를 필요로 하기 때문에 웨이모(미국), 바이두(중국) 등 글로벌 선도 기업들은 실증 거리를 확대하며 학습데이터 구축량을 늘리는 데 주력하고 있습니다.

그러나 국내의 경우 자율주행 학습데이터를 각 기업·기관이 개별적으로 구축하고 있고, 차종에 따라 센서 위치 등이 다를 경우 학습데이터를 상호 공유하기 어려워 데이터 부족 문제가 지속되어 왔습니다. 이는 인공지능(AI)의 성능이 데이터의 양에 비례하는 E2E 방식의 기술개발에 있어 병목점으로 작용해 왔는데, 과기정통부는 이러한 상황에서 학습데이터의 공유·활용을 극대화하고 글로벌 기술 동향 변화에 신속히 대응하기 위해 이번 가이드라인을 마련하였습니다.

### 2. 가이드라인의 주요 내용

본 가이드라인은 자율주행 E2E AI 학습에 필수적인 데이터 전(全) 주기를 포괄하며, 국내 연구진의 연구 편의성을 높이는 데 집중하였습니다. 구체적으로 ① 데이터 수집·가공·정합·보정·라벨링 등 자율주행 E2E 학습데이터 구축 절차를 정의하는 것에서 시작하여, ② 센서 구성과 저장 포맷 등 수집 시스템 요구사항 및 수집된 원시데이터의 검증 방법, ③ 시나리오 선별 등 데이터 가공 방법, ④ 위치 보정·공간 정

합 등 정합·보정 방법, ⑤ 라벨링 항목과 이를 통해 생성되는 학습데이터 셋 규격 등 라벨링 방법을 정리한 후, ⑥ 이를 기반으로 구축한 학습데이터 사례를 제시하고 있습니다.

본 가이드라인은 범부처 자율주행 R&D 사업인 「자율주행기술개발혁신사업」(’21년~’27년, 과기정통부·산업부·국토부·경찰청 협업)의 과제를 통해 개발한 과기정통부 핵심 성과 중 하나로, 정보통신기획평가원(IITP), 자율주행기술개발혁신사업단(KADIF), 한국전자통신연구원(ETRI)이 과제 목표를 상향 조정하면서 한국전자통신연구원(ETRI) 주도로 개발하였습니다. 또한 과기정통부는 초안 수립(’25년 말) 이후에도 한국 ITS 학회 특별세션(’26. 4. 23.), 자율주행 산·학·연 간담회(’26. 6. 9.) 등을 통해 업계와 전문가 의견을 지속적으로 수렴하였으며, 국토교통부와 협력하여 도시 단위로 대규모 E2E AI 학습데이터 구축을 지원하는 자율주행 실증도시 등에 실제 활용하면서 지속 보완해 나갈 계획이라고 밝혔습니다. 본 가이드라인은 자율주행기술개발혁신사업단 홈페이지(www.kadif.kr)에서 누구나 자유롭게 열람할 수 있습니다.

### 3. 시사점

**(산·학·연 협업을 통한 공동 데이터 구축 모델의 확산 가능성)** 이번 가이드라인은 각 기업·기관이 개별적으로 구축해 온 학습데이터의 상호 공유·활용을 위한 공통의 데이터 레퍼런스를 국내 최초로 정립하였다는 데 의의가 있습니다. 데이터의 양이 곧 기술 경쟁력으로 직결되는 AI 분야의 특성상, 이처럼 정부가 표준화된 규격을 제시하고 산·학·연이 데이터를 공동으로 구축·공유하는 협업 모델은 자율주행 분야를 넘어 향후 다양한 AI 산업 영역으로 확대될 수 있을 것으로 보입니다. 따라서 관련 기업으로는 이러한 공동 데이터 구축·활용 체계에 적극적으로 참여하여 실질적인 사업 기회를 모색하는 자세가 필요해 보입니다.

**(E2E 방식의 적용 범위에 대한 유의점)** 다만 ‘E2E’라는 용어는 자율주행 분야에서는 센서 입력부터 제어 명령까지를 하나의 AI 모델로 구현하는 방식을 의미하나, 보다 일반적으로는 입력값과 최종 라벨(결과)이 곧바로 연결된 구조, 즉 중간 단계의 판단 과정을 별도로 거치지 않고 입력에서 최종 결과를 바로 도출하는 구조를 가리키는 의미로도 사용됩니다. 이러한 구조는 사전에 정의되지 않은 상황에 유연하게 대응할 수 있다는 장점이 있으나, 결론에 이르는 중간 과정에 대한 설명이 어렵다는 한계도 있습니다. 이에 따라 금융·의료·법률 등 최종 결론에 대한 설명가능성(explainability)이 요구되는 분야에서는 순수한 E2E 방식을 그대로 적용하기에는 제한이 있을 수 있으며, 이러한 분야에서는 E2E 방식에 중간 단계에서의 전문가 개입을 결합한 하이브리드 방식을 활용하는 것이 보다 안전한 접근이 될 수 있다는 점에 유의할 필요가 있습니다.

## About Shin & Kim’s ICT Group

법무법인(유) 세종 ICT그룹은 ICT 분야의 독보적인 전문성과 인적 네트워크를 보유하고 있으며, 고객들로부터 최근 수년간 가장 높은 평가를 받고 있습니다. 방송과 통신, 개인정보, 인터넷 IT 분야에서 축적된 역량을 바탕으로 방송·통신·ICT 규제 동향 파악 및 대관, 법제개선·입법컨설팅, 규제영향력 분석과 기업의 전략 수립 등에 대한 종합적인 법률자문을 제공하고 있습니다. 침해사고 대응 등과 관련하여서도 다양한 업무경험과 전문성을 보유하고 있으므로, 보다 전문적인 내용이나 궁금하신 사항이 있으면 언제든지 연락 주시기 바랍니다.

## 관련구성원

## 강도현

고문

02-316-2884

dhkang@shinkim.com

## 노진홍

변호사

02-316-1639

jhnoh@shinkim.com

## 강지현

변호사

02-316-1518

jhykang@shinkim.com

## 강신욱

대표변호사

02-316-4059

sokang@shinkim.com

## 이지은

선임연구위원

02-316-1720

jeunlee@shinkim.com

## 정소정

변호사

02-316-1877

sjjeong@shinkim.com