



## 농촌진흥청, '농생명 특화 거대언어모델' 개발 방법 공유

- 27일 '농생명 특화 거대언어모델 개발 방법 공유 연수회' 열어
- 학습 데이터 구축부터 농생명 특화 거대언어모델 시연까지 기술 전수
- 개발 전 과정 담은 실무 지침서도 배부

※ 일시 및 장소: 2026. 10. 27.(화) 14:00, 농업생명자원부 1층 대회의실

농촌진흥청(청장 이승돈)은 10월 27일 국립농업과학원 농업생명자원부 대회의실에서 '농생명 특화 거대언어모델(LLM)\* 개발 방법 공유 연수회'를 열었다.

\*거대언어모델(LLM): 대용량 인간 언어를 이해하고 생성할 수 있도록 훈련된 인공지능 모형

농촌진흥청은 농생명 논문과 국가 연구데이터를 학습시킨 '농생명 특화 거대언어모델'을 개발했다. 이 과정에서 학습 데이터 구축·정제, 모델 정밀 조정(파인튜닝), 서비스 적용 등 핵심 기술을 확보했다.

이 자리는 최근 분야별로 특화된 소형 거대언어모델 요구가 급증하고 있는 점을 고려해 농촌진흥청이 농생명 특화 거대언어모델을 개발하며 쌓은 기술과 방법을 학회, 연구자들과 나누기 위해 마련됐다.

우리나라 맞춤형 농생명 특화 거대언어모델 개발을 위해 연구 정보를 공유한 국내 5개 학회 관계자를 비롯해 관계 기관에서 자체 인공지능을 개발 중인 부서 담당자 등 50여 명이 참석했다.

이날 행사는 △자료수집, 데이터베이스 품질 검토 등 훈련 데이터 구축 △농생명 특화 거대언어모델 개발 △시연 등 학습데이터 구축부터 서비스 적용까지 실무자 경험을 토대로 실제 개발자가 참고할 수 있도록 진행됐다.

이 중에는 문헌 데이터베이스 구축 시 가장 문제가 되는 한자 혼용 논문의 글자 인식, 정형·비정형 데이터 전처리 등 농촌진흥청 국립농업과학원만의 기술도 상당수 포함돼 있다.

이날 농생명 특화 거대언어모델 개발에 실제 적용한 방법을 담은 실무 지침서 ‘농생명 특화 거대언어모델 구축 기술 보급서’도 공개됐다. 이 책자는 도 농업기술원과 시군 농업기술센터 등에 배부될 예정이다.

농촌진흥청 슈퍼컴퓨팅센터 이태호 센터장은 “‘농생명 특화 거대언어모델’을 개발하며 얻은 기술과 방법을 공유해 연구개발(R&D) 기반이 약한 농업계 연구자들이 특화 거대언어모델을 만들 수 있도록 돕겠다.”라며, “이를 통해 농업계 전반에 특화 거대언어모델 구축과 활용 저변이 넓어지고, 협력관계도 다양하게 구축될 것으로 기대한다.”라고 말했다.

붙임. ‘농생명 특화 거대언어모델’ 개발 방법 공유 공동연수회

|       |                                                                                                                                                                      |     |     |                     |
|-------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----|-----|---------------------|
| 담당 부서 | 국립농업과학원                                                                                                                                                              | 책임자 | 센터장 | 이태 호 (063-238-4558) |
|       | 슈퍼컴퓨팅센터                                                                                                                                                              | 담당자 | 연구사 | 오재 현 (063-238-4766) |
|       | <div><div>농촌진흥청에서 연구·개발한</div><div>농업의 모든 것</div><div></div><div>농사호</div></div> |     |     |                     |

## 붙임

## ‘농생명 특화 거대언어모델’ 개발 방법 공유 공동연수회

- 일시 : 2026. 10. 27.(화) 14:00~16:20 (등록 13:30~)
- 장소 : 국립농업과학원 농업생명자원부 1층 대회의실(전북 전주)
- 참석 : 국립농업과학원장, 농업생명자원부장, 연구정보 공유 협약 5개 학회 관계자, 청 내 GPU 공동활용 참여 부서 담당자, 관계 부서 직원 등 40명 내외
- 주요 내용 : 학습데이터 구축부터 서비스 적용까지 단계별 실무 공유

| 시간          | 일정                                     | 비고       |
|-------------|----------------------------------------|----------|
| 14:00~14:10 | 개회 및 초청 인사 소개                          | 사회       |
| 14:10~14:20 | 인사말씀                                   | 국립농업과학원장 |
| 14:20~14:30 | 단체사진 촬영                                |          |
| 14:30~15:00 | 훈련데이터 구축<br>(자료수집, DB 구축 및 품질 검토 등)    | 윤혜진 연구사  |
| 15:00~15:30 | 농생명 특화 LLM 개발<br>(기반 모델 선발, 훈련 및 평가 등) | 이현아 박사   |
| 15:30~15:50 | 시연                                     | 오재현 연구사  |
| 15:50~16:10 | 고도화 반영을 위한 설문조사                        | 오재현 연구사  |
| 16:10~16:20 | 질의응답 및 폐회                              | 슈퍼컴퓨팅센터장 |

### □ 기술서

- 농생명 한국어 자료의 확보·정제, 학습데이터 구축, 모델 비교·학습, 검색 결합과 운영까지 ‘농생명 특화 거대언어모델’ 개발에 실제 적용된 방법론을 단계별로 정리한 실무 지침서
- 구성

|                 |                                    |
|-----------------|------------------------------------|
| PART 01 개요      | 모델 소개, 사용설명서                       |
| PART 02 데이터 준비  | 데이터 수집, 전처리, 품질검수, 증화·분할           |
| PART 03 모델 학습   | 추가 사전학습, 지도 미세조정, 성능평가             |
| PART 04 RAG 구축  | 문서 청킹, 임베딩, 벡터 DB 구축, 검색·답변 구현     |
| PART 05 에이전트 구현 | LangChain 연계, LangGraph 구현, 서비스 연계 |
| 부록              | 참고문헌, 용어설명                         |