
	<h1>보도자료</h1>	 한국과학기술정보연구원 <small>Korea Institute of Science and Technology Information</small>
배포 즉시 보도 가능합니다.		
대전(본원): 대외협력실 이종성 042-869-0976 / 이해준 0676 / 손영주 0997 문의: 연구데이터공유센터 이승복 책임연구원(042-869-0653)		
배포번호 : 2021-26 배포일자 : 2021.09.10.(금)	매수 : 보도자료 5매 (첨부자료 포함)	배포처 : 대외협력실

KISTI, 2021년 연구데이터·AI 분석활용 경진대회 개최

- 국가연구데이터플랫폼 DataON 저변 확대 도모 -

한국과학기술정보연구원(원장 김재수, 이하 KISTI)이 연구데이터 공유와 활용 활성화 및 DNA 저변 확대를 위해 과학기술정보통신부와 국가과학기술연구회가 주최하고 KISTI가 주관하는 ‘2021년 연구데이터·AI 분석활용 경진대회’를 개최한다.

올해로 2회를 맞는 본 경진대회는 ‘DataON* 또는 개인 보유 연구데이터를 사용한 분석·활용 사례 발굴 및 인공지능 모델 개발’을 주제로 일반적인 연구데이터의 분석·활용 사례를 보여주는 연구데이터 부문과 기계학습이 가능한 데이터셋을 기반으로 최적 성능을 보이는 인공지능 모델을 개발하는 인공지능 부문으로 진행된다.

* DataON: KISTI가 운영하는 국가연구데이터플랫폼서비스로서 연구데이터를 체계적으로 공유·관리하고 검색·분석·활용을 지원한다. 또한 커뮤니티 연구자들끼리 데이터를 공유하고 공동 활용해 국가 R&D 투자 효율성을 높이고 연구 생산성 향상에 기여하는 것을 목적으로 운영하고 있다.

분석환경은 KISTI에서 제공하며 연구데이터 부문은 DataON에서 모든 과정을 수행해야 한다. 데이터는 DataON 데이터 및 공개된 데이터와 개인 보유데이터를 활용할 수 있다.

이번 경진대회의 인공지능 부문에서는 작년 KISTI가 디지털 뉴딜 사업을 통해 구축한 과학기술 기계학습 데이터 5종 4백 59만여 건(첨부2 표 참조)을 공개하고 활용을 지원한다. 해당 데이터셋은 KISTI AI Data Archive(<https://aida.kisti.re.kr>)에서 연구 목적에 한해 이용 동의를 얻은 후 무료로 접근 가능하며, 향후 과학기술 논문 전문 기반 사전학습 언어모델도 10월 공개할 예정이다.

이번 경진대회는 1차 서류 평가와 최종 발표평가로 진행되며 연구 자료의 분석·활용과 인공지능 분석 등에 관심 있는 국민 누구나 개인 또는 4인 이하의 팀으로 참여할 수 있다. 서류 접수는 9월 23일(목)부터 30일(목)까지이며 DataON 공식 홈페이지(dataon.kisti.re.kr)를 통해 접수할 수 있다.

시상은 ▲최우수상(부문별 1명(팀)) ▲우수상(부문별 2명(팀)) ▲장려상(부문별 2명(팀))으로 진행된다. 최우수상은 각 상금 200만원과 과학기술정보통신부장관상, 우수상과 장려상은 각각 상금 150만원, 100만원과 국가과학기술연구회이사장상과 KISTI원장상이 수여된다.

KISTI 김재수 원장은 “연구데이터에 관심 있는 모든 분이 마음껏 기량을 펼치고 숨은 인재들이 많이 발굴되길 바란다” 며 많은 참여를 당부했다. 또한 “이런 기회를 통해 국가연구데이터플랫폼(DataON)이 더 많이 알려지고, 국가 차원의 연구데이터 분석 및 활용에 기여할 수 있기를 바란다” 고 밝혔다.

dataon.kisti.re.kr



과학기술정보통신부
Ministry of Science and ICT



nst 국가과학기술연구회
National Research Council of Science & Technology



KISTI 한국과학기술정보연구원
Korea Institute of Science and Technology Information

2021 연구데이터-AI 분석활용 경진대회



연구데이터가 만드는
새로운 연구패러다임

서류접수기간

9.23 목 - 9.30 목

공모내용

DataON 또는 개인 보유 연구데이터를 사용하여
분석활용 사례 발굴 및 인공지능 모델 개발

참가자격

연구데이터 공유·활용 및 인공지능 분석 등에
관심있는 분은 누구나 참여 가능

※ 개인 또는 팀(최대 4명)으로 참여 가능

공모부문

연구데이터 부문

일반적인 연구데이터의 분석활용 사례 개발

인공지능 부문

기계학습이 가능한 데이터셋을 기반으로
최적 성능을 보이는 인공지능 모델 개발

응모방법

온라인 접수

dataon.kisti.re.kr

대회일정 모든 일정은 상황에 따라 비대면으로 진행 가능



대회공고
9.8 (수) - 9.29 (수)



서류접수
9.23 (목) - 9.30 (목)



워크숍
10.5 (수)



서류평가 발표
10.6 (수)



사례 및 모델 개발
10.7 (목) - 11.18 (목)



발표자료 제출
11.18 (목)



최종발표평가
11.24 (수)



시상식
12.2 (목)

시상내역

총상금 **1400만원** / 총 10개 팀 선발

구분	연구데이터 부문	인공지능 부문	상금 및 부상
최우수상	1명(팀)	1명(팀)	상금 200만원 과학기술정보통신부장관상
우수상	2명(팀)	2명(팀)	각 상금 150만원 국가과학기술연구회이사장상
장려상	2명(팀)	2명(팀)	각 상금 100만원 한국과학기술정보연구원장상

문의처

2021 연구데이터·AI 분석활용 경진대회 사무국
T. 042-224-7175 E. interplan7@naver.com

첨부 2

KISTI 과학기술 기계학습 데이터 공유활용 시범서비스

□ 개요

- AI 기술 경쟁력 제고 및 저변 확대를 위해 과학기술 기계학습 데이터 분석·활용 인프라 구축·제공
- 과학기술 기계학습 데이터 5종(8월)과 과학기술 논문 전문 기반 사전학습 언어모델(10월)을 우선 공개
- KISTI KISTI AI Data Archive(aida.kisti.re.kr)에서 공개하며, KISTI ScienceOn(scienceon.kisti.re.kr)과 DataOn(dataon.kisti.re.kr)을 통해서도 검색과 접근이 가능하도록 제공 예정
- 향후 분야별 기계학습 데이터, 사전학습 모델, 관련 코드, AI 모델 구축 환경 등 과학기술 AI 데이터 분석·활용 인프라 구축 및 Data/AI 기반 문제 해결 체계 추진 예정



<서비스 구성도>

□ 공개 예정 과학기술 기계학습 데이터 소개

- '20년 과학기술 기계학습 데이터 구축 사업을 통해 구축된 데이터 5종과 대전시 도로영상 객체인식 데이터셋을 공개
 - 해당 데이터셋은 연구 목적에 한해 이용 동의 후 무료로 접근 가능

구축 데이터		건수	구축 내용
국내 논문	전문 텍스트	504,000	국내논문 PDF를 기계학습이 가능한 텍스트 형태로 구축
	질의응답 셋	283,000	국내논문에서 질의와 정답 쌍 구축
	문장의미태깅	155,000	국내논문의 연구목적, 연구방법, 연구결과 등 문장 의미 태깅
보고서	표/그림 설명	3,419,000	국가R&D보고서 원문의 표/그림에 대한 설명 텍스트 구축
	기관식별	233,000	국내 논문/보고서에서 저자의 소속기관명 식별데이터 구축
	대전시 도로영상 객체인식	-	KISTI가 자체적으로 수집한 도로영상 비디오를 활용한 영상객체 인식용 학습데이터셋
	합계	4,594,000	

□ 과학기술 기계학습 데이터 공유·활용 시범서비스 주요 기능

- 과학기술 기계학습 데이터 등록, 관리, 검색, 조회
- 데이터 활용 신청 및 Open API 제공
- 과학기술 논문 전문 기반 사전학습 언어모델 제공
- 데이터 분석·활용을 위한 DataOn 등 인프라 연계 기능 제공