

인공지능 학습용 데이터 구축 공모 설명회 개최 계획(안)

□ 목 적

- 인공지능 학습용 데이터 구축 사업의 사업 예산, 추진방향 사전 안내 및 관심있는 기관·기업의 참여 유도
- 사업참여 희망 기관·기업 등과 사업 계획, 추진방식, 주요일정 등 공유를 위한 상세 설명회 개최

□ 개 요

- (일시/장소) '19.2.11(월) 14:00~17:00 / 한국정보화진흥원 서울청사 지하1층
- (참석자) 인공지능(AI) 사업에 관심있는 민간 기업 및 유관기관 관계자
- (주최/주관) 한국정보화진흥원

□ 프로그램 계획(안)

시 간		내 용	비 고
14:00~14:05	5'	참석자 등록	
14:05 ~ 14:30	25'	인공지능 학습용 데이터 구축 추진방안	사회자
14:30 ~ 15:30	60'	과제별(10개) 요구사항 안내	담당자
15:30 ~ 16:00	30'	제안요청 사항, 사업신청서 및 계획서 작성방법 추진절차 등 안내	담당자
16:00 ~ 16:50	50'	질의·응답	사회자
16:50~17:00	10'	종합정리 및 폐회	

※ 질의응답 일정은 세부 사업에 따라 변동될 수 있습니다.

붙임 1 인공지능 학습용 데이터 구축

□ 사업 개요

- 지능정보산업 활성화를 위해 시급성과 유망성이 높은 분야를 발굴하여 산업적으로 활용가치가 높은 인공지능 학습용 데이터셋 구축 및 개방

□ 추진 전략

- 기업, 전문가 대상 인공지능 데이터 수요조사분석 추진('18.11월)
 - * 산·학·연 전문가 온라인 의견 수렴
- 수요 기반의 인공지능 데이터 중장기 로드맵(5개년) 수립('18.12월)

□ 세부 추진 과제

- (AI 데이터 구축) AI 기술·산업적으로 활용가치가 높은 공공·민간 데이터를 기계가 이해할 수 있는 AI 학습용 데이터 셋 구축·개방
 - CCTV 영상, 대화 음성, 사물 이미지 등 AI 서비스 개발에 공통적으로 활용될 수 있는 범용 데이터셋 구축 및 개방
 - 산업분야의 경우 시장·경제효과가 큰 헬스케어, 제조 등 분야별 대규모 데이터셋을 매년 10개 내외로 구축
 - 공공부문의 경우 행안부 주관 범정부 차원의 공공 데이터 전수 조사('18.8월)를 바탕으로 활용가치가 높은 데이터셋을 매년 5개 내외로 구축

□ 주요 일정

- '19년 3월 : '19년 AI 데이터 허브 사업공고
- '19년 4월 : 선정평가 및 수행기관 선정, 협약체결 등
- '19년 7월 : 중간보고 및 1차 데이터 공개
- '19년 11월 : 과제 완료 및 추진결과 보고 등

붙임 2

2019 인공지능 학습용 데이터 추진내용

- (AI 데이터 구축) 다양한 AI 학습용 데이터 셋에 대한 수요조사 및 전문가 평가를 통해 '19년 추진과제(10개) 선정 및 추진

구분	구 축 내 용	서비스 예시
① 번역 말뭉치	○ 한국어 인공지능 번역 기술개발 및 성능강화를 위한 한국어-영어 병렬 말뭉치 데이터 80만 문장 구축	한국어-영어 AI번역시스템
② 사물 이미지	○ 한국형 객체, 장소, 상황 인지기술 개발 및 성능강화를 위한 사물/거리/간판/심볼 이미지 데이터 360만장 구축	국내 사물인식 및 안내 서비스
③ 글자체 이미지	○ 한글 광학글자인식(OCR) 성능개선을 위한 한글 글자체(손글씨 및 인쇄체) 이미지 500만장 구축	지동 한글 인식 서비스
④ 인도 보행 영상	○ 시각장애인/전동휠체어 등의 보행지원기술 개발을 위한 국내 인도·횡단보도 보행 영상 및 인도위 객체(사람, 자전거, 가로수, 펜스 등) 라벨링 데이터 500시간 구축	장애인 길안내 서비스, 자율전동 휠체어
⑤ 융합 영상	○ 감성인식 AI개발을 위해 동영상에서 인물의 표정, 음성, 발화 내용, 상황 등의 감정이 포함된 멀티모달 영상 데이터 50시간 ※ 동영상 원천데이터는 저작권을 확보한 드라마, UCC 등의 영상 데이터 활용	인간의 감정, 상황 이해 로봇
⑥ 사람 동작 영상	○ 사람의 동작·자세·행동 인식기술 개발을 위해 다양한 조건에서 사람 동작 영상 데이터 50만 클립 구축	가정, 공공시설 등의 서비스 로봇
⑦ 안면 이미지	○ 다양한 각도, 조도 등의 환경 하에서 안면인식·식별 성능강화를 위한 한국인 얼굴 이미지 데이터 1,800만장 구축(600명)	한국인 얼굴 인식 AI시스템 범주자 탐지
⑧ 위험물 이미지	○ 위험물·도구 자동판별 기술개발 및 성능개선을 위한 위험물, 범죄 도구, 반입금지물품 등의 X-ray 이미지 42만장 구축	공항,항만,철도,주요 시설 AI보안 검사 시스템
⑨ 질병 진단 이미지	○ 국내에서 유병률, 중증도 등이 높은 주요 질환관련 진단 이미지(X-ray, MRI, 초음파 영상 등) 및 진단결과 데이터 1만장 구축	주요질병 자동검진 AI의사
⑩ 이상 행동 영상	○ 주야간 적용가능한 이상행동 지능형 탐지기술개발 및 성능 강화를 위한 이기종(가시광선, 적외선) 영상 데이터 300시간 구축	지능형 CCTV, AI안전감시 시스템