

세부사업	산업AI용데이터전처리자동화기술개발(R&D)					
과제명	산업AI용 데이터 전처리 자동화 기술개발			안전관리형 과제		-
				보안 과제		-
산업기술분류	대분류	정보통신	중분류	소프트웨어	소분류	SW솔루션
개요 및 필요성	<ul style="list-style-type: none"> ○ (개요) 산업데이터를 인공지능이 학습 가능 하도록 가공하는 산업데이터 전처리 자동화 플랫폼 개발 및 기업이 참조 가능한 공종별 라이브러리 개발 ○ (필요성) 데이터 전처리는 AI 모델 성능에 매우 큰 영향을 미치나, 중소·중견기업은 기술력·인력 등의 부족으로 산업데이터 전처리에 대한 어려움 해결 필요 					
과제목표	<ul style="list-style-type: none"> ○ (최종목표) 전처리 자동화 플랫폼 및 5대 공종 분야(안전, 품질, 설비, 환경·에너지, 공정·물류) 라이브러리 개발 ○ (지원대상) 산업데이터 및 전처리 가공 기술을 보유하고 AI와 연계·활용하고자 하는 (주관연구개발기관) 비영리기관 / (공동연구개발기관) 공급기업 및 중소·중견(수요) 등 컨소시엄 ○ (필수지표) <ul style="list-style-type: none"> - 데이터 전처리 자동화 플랫폼 개발 1식 - 전처리 자동화 라이브러리 개발 30종 이상 ○ (자율지표) <ul style="list-style-type: none"> - 연차별 전처리 라이브러리 개발 종수 - 전처리 시간단축률(기존 수작업 전처리 대비 자동화 플랫폼 도입 후 절감된 시간의 비율) - 전처리 자동화율(전체 전처리 단계 대비 자동화된 전처리 단계의 비율) - 전처리 수행 전후 데이터 품질 향상 요소 - 실증 기업 수 및 활용 기업 수 					
과제내용	<ul style="list-style-type: none"> ○ 데이터 전처리 자동화 플랫폼 개발 : 산업데이터를 효율적으로 처리하는 자동화 플랫폼 개발 및 사용자 친화적 인터페이스 제공 <ul style="list-style-type: none"> - 업종별 Pain Point가 발생하는 핵심 공종별 AI 모델 선정 및 데이터 전처리 방법론 수립 - 로코드·노코드 기반 데이터 전처리 자동화를 위한 전처리 수행 기능 등 전처리 필요 기능 및 전처리 자동화 알고리즘 설계* * 데이터 특성을 자동으로 분석하여 최적의 전처리 방식 및 라이브러리 제안 - 데이터 전처리 방법 수행 순서를 정의하고, 쉽게 순서 및 기능을 변경할 수 있는 프로세스 모델러 개발 - 프로세스 모델러와 수행 기능 등이 자동으로 수행 시 모니터링 및 관리 기능 개발 - 데이터 전처리 프로세스 모델러, 수행 기능, 자동화 기술 및 					

	<p>모니터링 기술을 종합한 산업데이터 전처리 자동화 솔루션 개발</p> <p>* 데이터 전처리 자동화 솔루션은 현업 종사자가 사용이 용이하도록 데이터 전처리 방법, 파라미터 및 순서 등의 변경 방법이 직관적인 설계</p> <p>○ 전처리 자동화 라이브러리 구축 : 전처리 자동화 플랫폼에서 사용될 5대 공종(①안전, ②품질, ③설비, ④환경·에너지, ⑤공정·물류)에 특화된 전처리 라이브러리 30종* 이상 개발</p> <p>* 산업 수요가 많은 공종 중심으로 구축하되 공종별 최소 5종 이상 구축, 수요기업과 연계하여 실제 산업데이터 활용하여 개발</p> <p>- Pain Point에 직간접적으로 영향을 미치는 산업데이터 조사, 既 확보 데이터 검증, 추가 필요 데이터 정의</p> <p>- Pain Point를 산업 데이터 기반의 문제로 정의하고, 이를 해결하기 위한 다양한 AI 모델 중 적합 모델 선정 및 데이터 활용 테스트</p> <p>- Pain Point를 해결하기 위한 반복적 AI 모델 선정 과정을 통해 전처리 모듈 개발 및 표준화</p> <p>- 선정한 AI 모델 수행을 위한 데이터 전처리 방법의 기능과 프로세스 등을 S/W 라이브러리로 개발</p> <p>- 전처리 라이브러리는 데이터 분류(정형, 음성, 영상 등)에 대한 균등한 분포를 고려하여 제시</p> <p>○ 산업 실증 및 확산지원체계 구축 : 전처리 자동화 플랫폼의 산업현장 실증을 통한 성능 검증 및 산업확산지원 체계 구축</p> <p>- AI 모델, 데이터 전처리 방법론, 프레임워크 등 솔루션과 신규 발생 산업데이터를 활용하여 Pain Point 해결 여부에 대한 실증 및 검증(성능평가)</p> <p>- 주관연구개발기관(대학/연구소/비영리법인)이 플랫폼 공개 시 수요기업이 최소한의 비용으로 활용할 수 있도록 관련 활용 방법 제시 및 지속가능 운영방안 제시</p>		
기대효과	○ (기대효과) 데이터 전처리 소요 시간 단축 및 비용절감, 산업 데이터 활용도 증대 등		
총수행기간	2025년 - 2028년 (4년) (1차년도 연구개발기간 : 9개월)	총 정부지원연구개발비* (*25년 정부지원연구개발비)	18,250백만원 이내 (3,200백만원 이내)
주관연구개발기관	<input checked="" type="checkbox"/> 대학 <input checked="" type="checkbox"/> 연구소 <input checked="" type="checkbox"/> 비영리법인 <input type="checkbox"/> 제한없음		
공동연구개발기관	<input type="checkbox"/> 대학 <input type="checkbox"/> 연구소 <input type="checkbox"/> 비영리법인 <input checked="" type="checkbox"/> 제한없음		

* 상기 정부지원연구개발비는 예산 현황 및 평가 결과에 따라 변동될 수 있음