

데이터 품질진단 및 개선 사업 기업 설명회

2025.7.16.(수) 오후 2시~4시

(주)페블러스

Chapter I

사업 개요

- ① 사업 배경 및 목적
- ② 사업 범위
- ③ 페블러스 솔루션
- ④ 사업 추진전략

데이터 품질 진단 및 향상 지원을 통해 지역 주력 산업 경쟁력을 강화하여 고품질 데이터 기업 육성

사업 배경

- ☞ 산업별 AI 대전환 가속화

 - 미래 모빌리티·로봇·헬스케어·반도체 등 지역 핵심산업은 'AI 효율'이 곧 경쟁력
 - 학습 가능한 AI-Ready Data 확보가 필수
- ☞ 데이터 품질의 현실

 - 오류·중복·결측·편향 → AI 모델 성능 및 효율 저하
- ☞ AI 규제 & 컴플라이언스

 - EU AI Act, 국내 AI 기본법(제정 추진) → 데이터 출처·품질·설명가능성 증명이 필수 조건으로 부상
- ☞ “AI-Ready Data가 기업의 리스크를 줄인다.”

 - 위험 및 신뢰의 문제 Risk & Trust
 - 부적격 데이터는 AI 모델의 편향 및 환각 유발
 - 규제기관의 “데이터 검증” 요구
 - 지속적인 변동의 문제 Continuous Change
 - 데이터 값·분포·컨텍스트가 상시 변동
 - 상시 모니터링 필수
 - 거버넌스 의무 Governance Mandate
 - 품질·메타데이터·투명성의 지표화 동향: EU AI Act, 국내 공공데이터 활용도 평가
 - AI-Ready 체계 없이는 사업·조달 모두 제동 예상

사업 목적

- ☞ 데이터 품질 진단·향상(증강·경량화)으로 AI-Ready Data 구축
- ☞ 품질·보안·윤리를 포괄한 데이터 거버넌스 체계 구축 지원
- ☞ 고품질 데이터 자산을 확보한 지역 주력 산업 ‘챔피언 기업’ 육성
 - 산업 경쟁력 + AI 활용도 동시 제고



대구 지역기업 데이터 품질 향상을 위한 데이터 활용 컨설팅 및 품질 진단/개선 수행

과업 목표

수요기업 모집

사전 컨설팅

데이터 품질진단

데이터 증강 및 경량화

후속지원

제안 범위

- 지원 대상
 - 대구·경북 소재 본사/연구소 보유 기업 중 AI 활용 잠재력이 높은 주력산업 6대 분야 (모빌리티·로봇·헬스케어·반도체·ABB·물산업)
- 데이터 유형
 - 정형(테이블) 및 비정형(이미지, 영상 등)
- 지원 범위
 - 컨설팅: 데이터 활용 및 AI-Readiness 컨설팅
 - 품질진단(20개사) : AI-Ready Data 단계별 진단 또는 페블러스 자체 솔루션 활용
 - 품질개선(10개사) : 합성데이터/경량화 패키지 적용
 - 후속 지원 : 거버넌스/모니터링 컨설팅

진단부터 AI 성능 개선까지, One-Stop 데이터 품질 관리 솔루션

AI 임베딩 기반 진단

- 정형/비정형 학습 데이터의 품질을 신경망 임베딩 기반 시각화로 분석
- 불균형, 중복, 노이즈 등 품질 이슈 자동 식별
- 기존 룰/통계 방식보다 깊이 있는 문제 탐지

데이터 품질 진단-처방-개선 분석 플로우

- 품질 문제를 '증상화'하고, 치료 방식(개선전략) 제시
- 직관적 시각화 리포트 + 정량 지표 + 개선 시뮬레이션
- 비전문 기업도 이해할 수 있는 실행 중심 보고서

스마트한 증강 & 경량화 - Data Bulk-Up / Diet

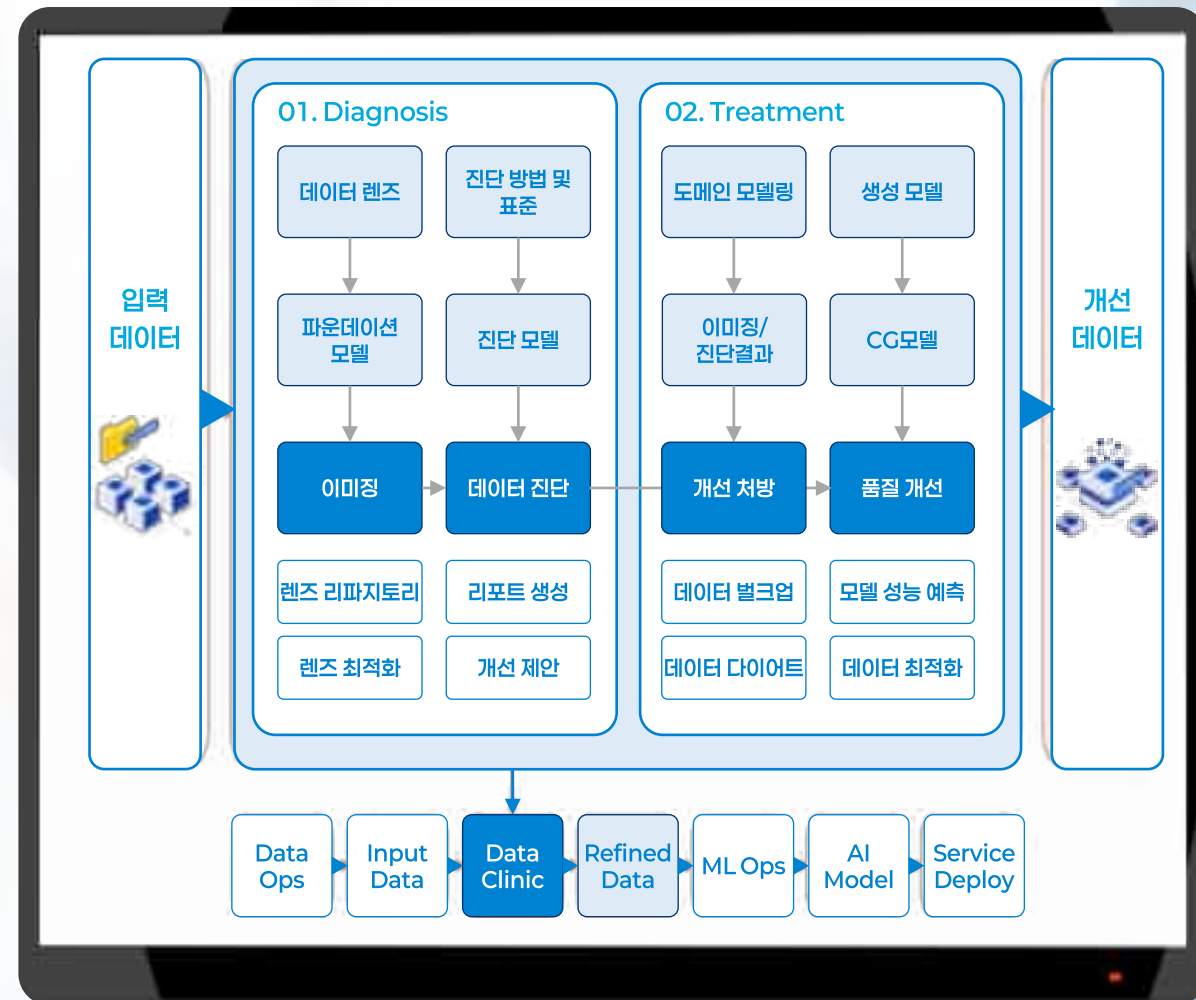
- 합성데이터 생성: 빈도/분포/유사성 기반 정밀 증강
- 데이터 경량화: 중복/저정보량 샘플 제거로 최적화
- 가명정보 처리 기준 완벽 반영

SaaS 연계 & 후속지원 가능

- 페블러스 개발 플랫폼 기반 클라우드 진단/모니터링
- 기업 맞춤형 서비스화 전략 지원 (데이터 상품화, API, SaaS 고도화)

검증된 실적 & 확장성

- 제조, 로봇, 의료 등 다양한 산업 경험 기반
- 자동화된 분석 파이프라인으로 빠르고 일관된 품질진단 가능



진단부터 AI 성능 개선까지, One-Stop 데이터 품질 관리 솔루션

데이터 품질 진단

고객의 AI 학습 데이터를 데이터 렌즈를 통해 임베딩하여 특정 벡터로 표현하고, 이를 과학적 지표와 가시화를 통해 품질을 측정하고 설명하는 진단 서비스를 제공합니다.

Step 1

데이터 이미징

데이터의 잠재력 발굴



데이터 이미징과 데이터 렌즈(Data Lens) 기술을 통해 대용량 데이터를 관찰 가능하고 측정 가능하게 변환합니다.

Step 2

품질 진단

정밀한 데이터 분석



이미징 결과에 대해서 데이터 커버리지, 밀도, 균질도, 이상치 등 데이터 품질의 다양한 지표를 측정합니다.

Step 3

데이터 리포트

품질에 대한 설명과 시각화



데이터 이미징과 진단의 결과를 구체적인 사례와 시각화와 함께 설명하고 품질 개선을 제안합니다.

Step 4

품질 개선

최적화된 품질 개선 서비스



데이터 백업(합성데이터 추가) 또는 데이터 다이어트(불륨 경량화) 같은 기법을 통해 데이터 품질을 개선합니다.

데이터 품질 개선

AI 모델 성능 향상을 위해 데이터셋을 진단하여 최적화된 합성데이터를 생성하거나 불필요한 데이터를 제거하는 서비스를 제공합니다.

자사 솔루션 및 진단 프레임워크를 활용하여 빠르고 객관적인 데이터 품질 진단 및 개선 수행

성공적인 사업을 위한 추진전략

추진전략	내용	기대효과
검증된 자사 방법론 및 솔루션 적용	자체 솔루션 적용	단기간 내 객관적 평가
AI-Ready Data 단계별 진단 방법론 적용	Gartner 평가 지표 활용 데이터 진단	AI-Readiness 현황 파악 및 방향성 제시
도메인별 Best Practice 내재화	제조/헬스케어/모빌리티/에너지 등 10여 건 수행 노하우	업종별 맞춤 컨설팅 및 품질 개선

데이터 품질관리 관련 유사 사업 수행 경험으로 인한 역량 보유

사업성과

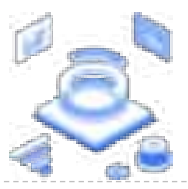
데이터바우처 사업 수행(2023, 2024)
AI 학습용 이미지 품질 개선 다수 경험



한화비전 데이터 품질 진단(2024)
지능형 CCTV 데이터 AI 활용 품질 진단



안전성평가연구소 데이터 경량화(2024)
신약개발 데이터 경량화



한국전력 데이터 품질 진단 및 컨설팅(2025)
전력 데이터 AI 활용 품질 컨설팅



Chapter II

수행 방법

- ① 수행 방안
- ② 수요기업 모집
- ③ 사전 컨설팅
- ④ 데이터 품질 진단
- ⑤ 데이터 품질 개선
- ⑥ 후속지원

5단계 핵심 과업 추진

수요기업 모집

데이터 활용 및 AI 전환 기업
수요 발굴을 위한 설명회 개최

산출물

과업별 수행일정표

수요기업 리스트

사전 컨설팅

수요기업 데이터 현황 및 활용
목적 분석 설문지 진단

오프라인/온라인 컨설팅

1일 최대 5개사

3주 안에 30개 기업 컨설팅 수행

서비스 방향 및 목표 수립

산출물

기업별 컨설팅 보고서

데이터 품질 진단

자체 개발 솔루션 활용

Gartner사의 AI-Ready
Data 평가지표 활용

산출물

기업별 품질진단 보고서

데이터 증강 및 경량화

품질진단 기반 컨설팅

데이터 증강 : 합성데이터 생성

데이터 경량화 : 중복 데이터
제거

산출물

데이터 증강/경량화
결과보고서(10개사)

후속지원

지속적인 데이터 품질
모니터링 지원

데이터 관리 거버넌스
구축 지원

산출물

결과보고서(해당기업)

희망기업 현황과 니즈를 파악하여 컨설팅 수요기업 확정

데이터 활용 및 AI 전환 기업 수요 발굴을 위한 "설명회" 개최

● 일정 및 장소 확정

- 2025년 7월 중
- 오프라인 개최

● 프로그램 구성

- AI 전환 트렌드 및 전망 소개
- AI-Ready Data 소개
- 맞춤형 데이터 컨설팅 서비스 소개
- Q&A

● 설명회 진행

- 데이터 품질 진단 전문가
- 합성데이터 전문가

● 희망기업 컨설팅 신청

- 희망기업의 데이터 활용 현황 및 니즈, 컨설팅 수요 등을 간략히 작성하여 신청
- 온라인/오프라인 개별 상담

● 컨설팅 수요 기업 선정

- 신청서 평가 및 개별 상담을 통해 컨설팅 대상 적합성 평가
- 설명회 참가 기업 중 데이터 활용 니즈가 높고 컨설팅 수요가 있는 기업 우선 선정

산업별 타겟 & 수요조사 절차

타겟 산업

- | | | |
|---|---|--|
| <input checked="" type="checkbox"/> 미래 모빌리티 | <input checked="" type="checkbox"/> 로봇 | <input checked="" type="checkbox"/> 헬스케어 |
| <input checked="" type="checkbox"/> 반도체 | <input checked="" type="checkbox"/> ABB | <input checked="" type="checkbox"/> 물 산업 |

Long-list 구축

- 산업 DB + 파트너 추천 : 30개사 이상

프리 설문

- 데이터 형태, 애로사항, 데이터 활용 목적 등

온라인/오프라인 상담

- 적합도 스코어링

Top 30사 선정

- 의향서/신청서 기반 & 컨설팅 예약

피드백

- 미선정 기업에 향후 가능 서비스 안내

주력산업 도메인 맞춤 접근 컨설팅 수행

현황 분석

수요기업의 데이터 현황 및 활용 목적 분석을 위한
사전 설문지 제공



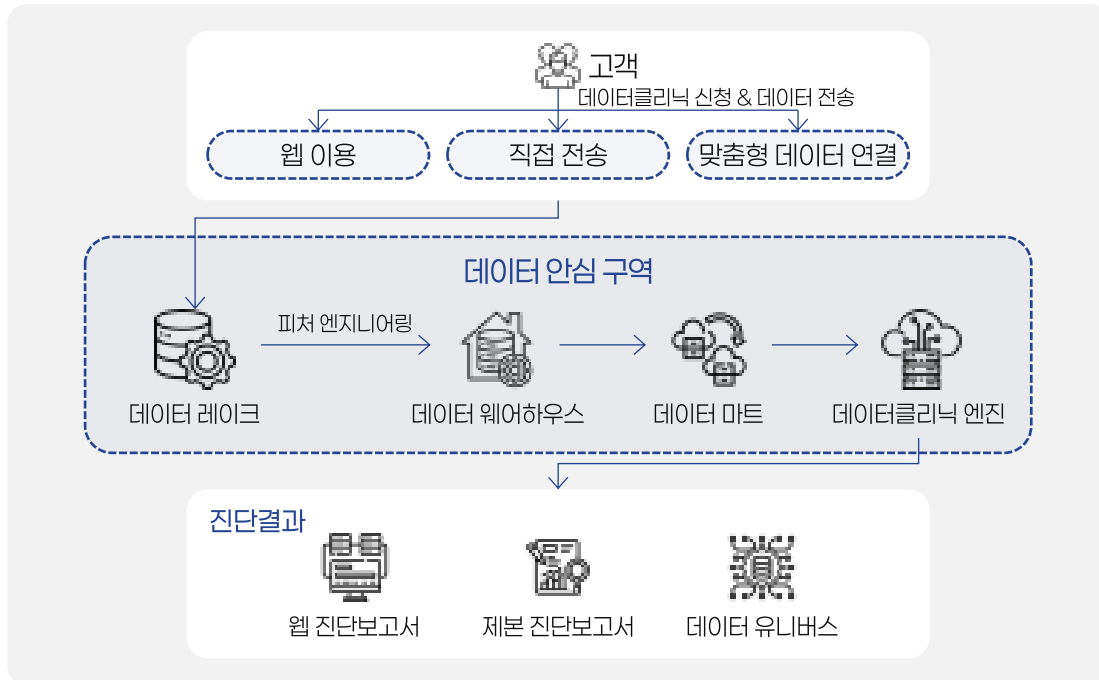
컨설팅 진행

- ☑ 기업 방문 또는 온라인 컨설팅 진행
 - 1일 최대 5개사
 - 수요기업과 일정 조율하여 3주 안에 30개 기업 컨설팅 진행
- ☑ 서비스 방향 설정
- ☑ 서비스 목표 수립
- ☑ 컨설팅 결과보고서 제공
- ☑ 개인정보 비식별화 가능성 논의

페블러스의 역량과 노하우를 결합하여 자체 개발한 데이터 품질 진단 및 개선 솔루션

데이터 품질 진단-처방-개선 분석 플로우

- ✓ 품질 문제를 '증상화'하고, **치료 방식(개선전략)** 제시
- ✓ 직관적 시각화 리포트 + 정량 지표 + 개선 시뮬레이션
- ✓ 비전문 기업도 이해할 수 있는 **실행 중심 보고서**



AI 임베딩 기반 진단

- ✓ 정형/비정형 학습 데이터의 품질을 **신경망 임베딩 기반 시각화**로 분석
- ✓ 불균형, 중복, 노이즈 등 **품질 이슈 자동 식별**
- ✓ 기존 룰/통계 방식보다 **깊이 있는 문제 탐지**



고품질 증강 데이터

저품질 데이터: 중복 데이터

Gartner사의 AI-Ready Data 평가지표 활용

- ✔ “리스크가 사람 책임 Skills-managed 에서 시스템 책임 System-based 으로 넘어갈 때 비로소 데이터는 진짜 AI-Ready가 된다.”
- ✔ AI-Ready 데이터를 위한 기능적 요구사항
 - 모드별 필수/선택 항목 점검
- ✔ 깊이 우선(Depth-first)
 - 자동화보다 “메타데이터 충실도”가 리스크 결정
- ✔ 메타데이터 그래프(Metadata Graph)
 - 구조·비정형 자산을 체이닝(chaining)하여 그래프/벡터 관리
 - 대표성·맥락 격차 상시 감시

영역	데이터 요구 사항	적용	참고사항
기본 적합성 Foundational Alignment			
Operational Maturity	일관성	일일-세일일 데이터 자산의 레코딩/업데이트 수량에 관계없	
Operational Maturity	가시성 지표	모든 주요 데이터 레코딩 접근 도메인(일-세일) 사용 사례 수를 모니터링	
Operational Maturity	내용 일관성	동일한 정보가 다른 형식/구조로 존재할 때 교차 참조로 접근 가능	
Operational Maturity	호스 수동 검증	설계 목적에 따라 원본 데이터가 정확히 수집되었는지 품질 목적으로 확인	
공정 사용 적합성 Qualified Usage (Fairness)			
Operational Maturity	백지 품질	일일 레코딩 100%는 데이터 접근(availability)과 무관성(immutable)의 특성 확인	
Operational Maturity	데이터 계보	데이터 출처에서 사용 사례까지의 전체 경로 및 분기(branch)를 추적	
Operational Maturity	권한 일관	공인된 시스템/제공자에서 데이터가 공유되는지 계보로 확인	
Operational Maturity	다양성 확보	같은 출처의 거래할 다양한 데이터 유형(텍스트, 이미지, 오디오, 비디오) 확보	
Operational Maturity	표준 성능	읽기-전송-편집 지연이 결과 계산에 미치는 영향 인지	
Operational Maturity	거분 접근 용이	원본 또는 고품질 보조 데이터 접근, 무관한 시스템은 차단	
Operational Maturity	다중계층 메타데이터 연결	내용-컴퓨트-문 등 분산 메타데이터를 일관하게 부여해 MAD 활용	
Operational Maturity	데이터-속성-속성 그래프 연계	연장 계보, 그래프 등 다양한 링크 확보-연결	
Operational Maturity	사용 사례 공유성	같은 문제를 다루는 여러 팀 시스템이 “공통 메타데이터”를 조직 전체에서 일관하게 적용	
모범 실행 적합성 Qualified Production (Fit for)			
Operational Maturity	백지 관리	데이터 스키마 변경을 문서화, 추적 가능 가능	
Operational Maturity	학습**관련 데이터 검증	보통-품질 무조건 확인 필요 시 변경 기록 및 패턴 분석	
Operational Maturity	변형 준수	기술-계보-공통-변형-확인 등 데이터 공유 제공	
Operational Maturity	규제 준수	전고 호스-기준-실제 데이터-기준-이전 사용물 문서화	
Operational Maturity	AI 기술-목적 적합성	모델-개발자 사용 목적에 대해 적합-효과-인도-변형-확인	
Operational Maturity	데이터 노출 위험	외부-기업-목적-거버넌스-요구-구분, 해당-데이터-공급-위험	
Operational Maturity	데이터 공유 링크	외부-제-데이터-링크-데이터-누출-감시-기록-분석, 해당-위험-대응-책	
Operational Maturity	콘텐츠 사용 가시성	유사-데이터-공통-구조-정보, 공유-데이터-추적	
적합-업선 대비 Augmented Qualified Data Readiness			
Operational Maturity	회용 테스트	데이터 변경이 기존 문제 해결을 이루는지 지속 검증	
Operational Maturity	연속 프로파일링 분석	실-무중단-프로파일링-데이터-변동-그래프-분석	
Operational Maturity	변화 인식 능력	변화-속도-가시-감시-변형-시-데이터-자료-모집-연속	
Operational Maturity	후원-데이터-일관성	모든-전송-데이터-원본-데이터-원본-데이터-원본-추적-가능	
Operational Maturity	파생 데이터 생성	Seed-데이터-원기-저장-모형/분석-데이터-원본-데이터-원본-추적-가능	

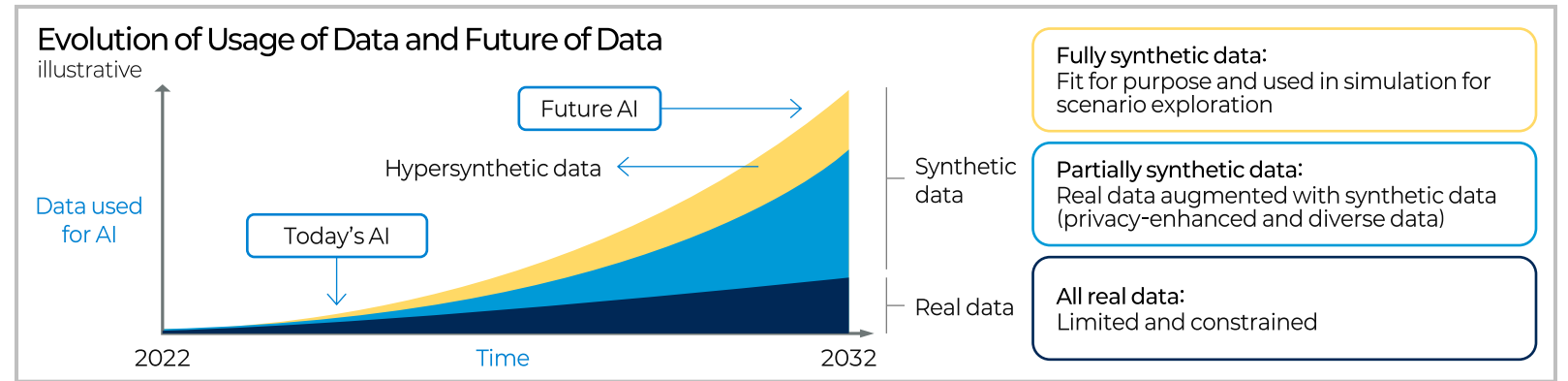
Gartner사의 AI-Ready Data 평가 항목 (재가공)

데이터 품질 진단기반 컨설팅

인공지능 학습을 위해

- (1) 데이터 수량이 부족한 경우,
- (2) 실제 데이터를 구할 수 없는 경우,
- (3) 다양한 환경에서의 데이터가 필요한 경우,

합성데이터를 제작합니다.

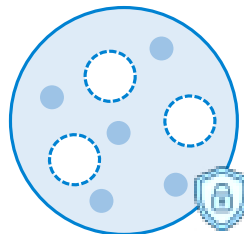


데이터 품질 개선 솔루션

○ 비어있는 데이터 영역 ● 중복 데이터 영역 ○ 실제 데이터 ○ 합성 데이터

데이터 레플리카

원본의 내용을 보호하면서도 통계적·분포적 특성이 유사한 합성데이터를 생성

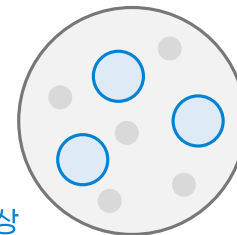


→ 개인정보 유출 리스크 예방

활용 예) 개인정보 보호를 위한 데이터 익명화

데이터 벌크업

데이터가 부족한 부분에 데이터셋에 적합한 정밀 타케팅 합성데이터를 생성하고 제공

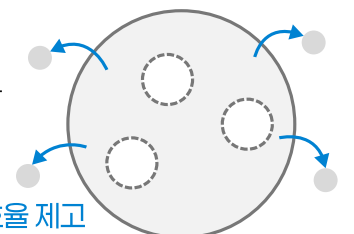


→ 데이터 수집 비용 절감, AI 성능 향상

활용 예) 예외 상황에 대한 합성데이터 생성

데이터 다이어트

AI 모델의 성능을 보장하는 선에서 데이터셋의 중복/유사 데이터를 제거하고 경량화



→ AI 학습 비용 절감, 개발 효율 제고

활용 예) 학습 과정에서의 데이터 경량화

데이터 증강 포트폴리오: 합성데이터 생성(3D 그래픽과 AI 생성 모델 하이브리드 기법)

드론의 인식 및 추적을 위한 3D 합성데이터



급식 섭취 모니터링을 위한 합성데이터



실제 이미지와 생성형 AI 모델을 활용하여 생성한 정밀 타기팅 합성데이터



차량 내부에서의 어린이 인식 등 특수환경에서의 인물 및 자세 인식을 위한 AI 학습용 합성데이터



3D 그래픽과 AI 생성 모델의 하이브리드 기법으로 국방 감시 체계 시를 위한 합성데이터 생성



데이터 품질 향상 지원 기업들에 대해 만족/불만족 사항을 파악하고 후속지원 실시

목표

품질 향상 기업의 **'만족 90점↑'**
지속 품질 유지율 95% 달성

- ☑ 만족도 조사
- ☑ 기업별 맞춤 후속 지원
- ☑ 성과 공유 & 확산
 - 사업 홍보
 - 지역 세미나 발표 등



Chapter III

수행사 소개 및 추진 일정

- ① 페블러스 소개
- ② 유사 사업 수행 실적
- ③ 수행 조직
- ④ 추진 일정

페블러스는 데이터 분야의 기술적 역량과 네트워크를 바탕으로 데이터 컨설팅 및 구축 분야에서 성공적인 이력을 쌓아 왔습니다. 2023년에 “인공지능 학습데이터 품질 진단 및 개선” SaaS 솔루션을 출시한 데이터 품질 전문 업체입니다.

기업 기본정보

항목	세부 내용	
기업명	주식회사 페블러스	
설립연월일	2021.11.10.	
대표자명	이주행	
업종(코드)	정보통신업(62010)	
대표전화	044-589-3824	
주요 생산품	데이터 품질 진단 및 개선 솔루션	
사업장 소재지	본사	대전광역시 유성구 대학로 99, 507호
	연구소	세종특별자치시 집현중앙7 6, B619호

주요 연혁

2021

- 06 - 중기부 예비창업패키지 선정 (최우수 종료)
- 09 - 대전창경 창업 아이디어 경진대회 1위
- 10 - 중소기업기술개발지원사업 TIPS 선정
- 11 - 법인 설립

2022

- 02 - 데이터/AI바우처 사업 공급기업 선정
- 04 - 현대자동차그룹 ZERO1NE 프로젝트 팀 선정
- 05 - 데이터바우처 사업 4건 매출 계약
- 06 - IBK 창공 혁신 창업기업 선정
- 11 - <도전! K-스타트업> 왕중왕전 장려상 수상

2023

- 05 - 신용보증기금 퍼스트펍권 선정
- 07 - 제1기 양재 AI 허브 멤버십 선정
- 07 - 하나은행xSK텔레콤 AI Lab 입주기업 선정
- 08 - IBK 1st Lab 선정 (금융합성데이터)
- 08 - 한국핀테크지원센터 D-테스트베드 선정
- 08 - 제조데이터 구매지원사업 선정
- 11 - 싱가포르 SWITCH 2023 전시 참여
- 11 - 한국경제 디지털이노베이션 정보기술분야 대상 수상

2024

- 01 - CES 2024 현대차 제로원 부스 전시 참여
- 02 - 한국특허전략개발원 특허·디자인 융합 IP-R&D 전략지원사업 선정
- 03 - TIPS 창업성장기술개발사업 성공 종료
- 03 - 현대자동차 ZERO1NE 크리에이터 선정
- 04 - IITP SW컴퓨팅산업원천기술개발사업 참여기관 선정
- 04 - 한국발명진흥회 2024년 제1차 IPL래 프로그램 선정
- 04 - DISS 2024 키노트 발표 및 부스 전시 참여
- 04 - AWS 스타트업 성장 프로그램 선정
- 07 - 초거대AI 확산 생태계 조성 사업 참여기관 선정
- 12 - 삼성E&A 프로젝트 수주

2025

- 01 - CES 2025 대전창조경제혁신센터 부스 전시 참여
- 02 - 초거대AI 확산 생태계 조성 사업 성공 종료
- 02 - 가트너 데이터·AI 추천기업 공식 등재
- 03 - MWC 4YFN 전시 참여
- 04 - 글로벌 기술사업화(아시아) 지원사업 선정
- 05 - LG전자 데이터 분석·시각화 프로젝트 수주
- 05 - AI EXPO 2025 부스 전시 참여
- 07 - 데이터 클리닉 솔루션 조달청 혁신제품 심사중(7월중 지정 예정)

페블러스는 품질 진단 및 개선 전문 기업입니다

● 합성데이터 품질 평가·개선 스타트업

- 국내 최초 Synthetic Data 제작 → 다수 기업 기술이전·시장 검증
- 창업자: 10년 CG·AI R&D, 합성데이터 분야 선구자

● 데이터 품질 진단 핵심 기능

- 데이터 부족·편향·중복·불균형·라벨 오류 자동 진단 (신경망 기반)
- One-Click Report : 품질 점수 · 원인 · 개선안 시각화

● 품질 개선 패키지

- ① 데이터 벌크업 - 희귀 케이스 합성
- ② 데이터 다이어트 - 중복 샘플 일괄 제거
- ③ 재현 데이터 - 오류 라벨 자동 검출·수정

● 데이터 커뮤니케이션 도구

- 대시보드 & PDF 리포트 자동 생성
- 이해관계자 맞춤형 시각화 템플릿 제공

● 확장 서비스

- 산업별 도메인 기반 맞춤형 컨설팅
- 클라우드형 데이터 관리·거버넌스 서비스



CEO | Co-Founder
이주행

Profile

- POSTECH 컴퓨터공학 박사 (Ph.D)
- ETRI 책임연구원 (CG, VR, Robotics, AI) 경력 23년 UST 교수
- 대전 비엔날레 초청 작가



COO | Co-Founder
이정원

Profile

- KAIST 바이오및뇌공학과 박사 (Ph.D)
- 서울대학교 전기공학부/의공학과 석사
- ETRI 책임연구원(AI) 경력 20년
- 의료 AI 개발
(개발 제품 서울대병원 설치 및 수출)

Gartner사의 데이터 품질 관리 보고서와 합성데이터 보고서에 각각 주요 기업으로 언급되는 성과를 이루었습니다.

생성 인공지능 적합 데이터(GenAI-Ready Data)를 지원하기 위한 비정형 데이터 관리 역량 개발, Gartner, 2025.

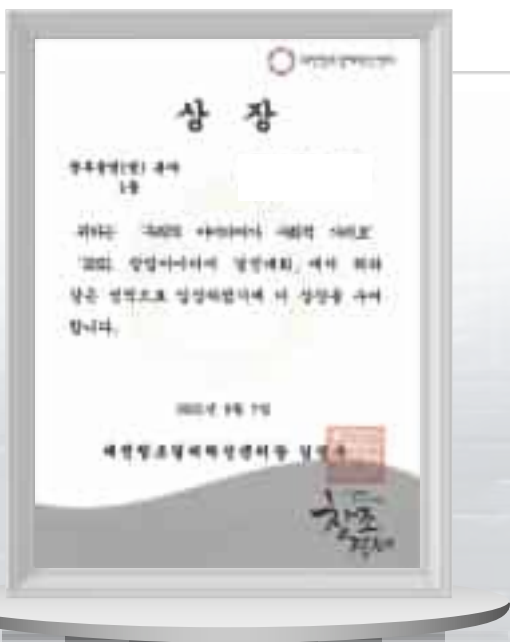
합성데이터 분야 스타트업들을 위한 기술 지형도, Gartner, 2025.



페블러스는 데이터 품질 평가 및 개선 분야에서 다양한 수상 내역을 보유하고 있습니다.

2021

대전창조경제혁신센터
창업 아이디어 경진대회 1위



2022

<도전! K-스타트업>
왕중왕전 장려상 수상



2023

TIPS-CNU
Startup IR 경진대회 2위



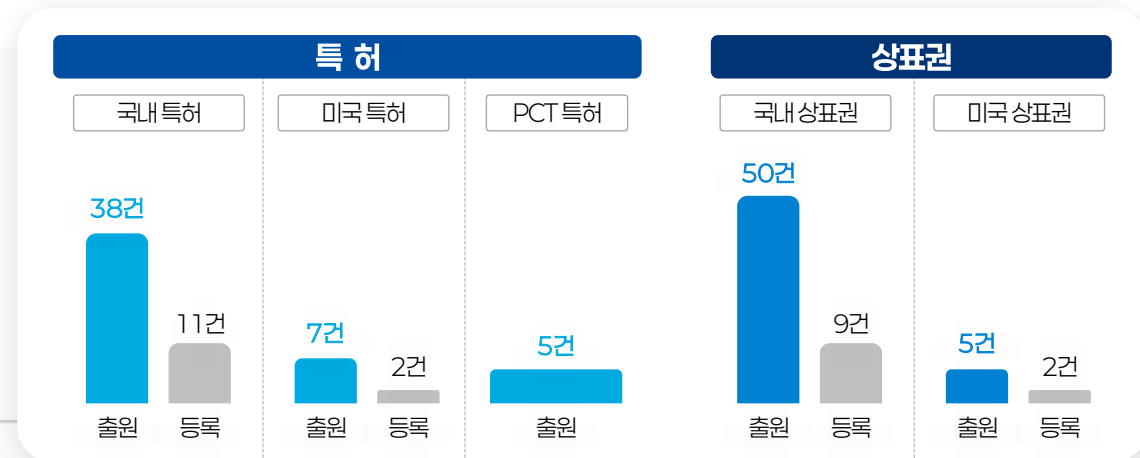
2023

한국경제 디지털이노베이션
정보기술분야 대상



다수의 국내외 지적재산권을 확보하고 있습니다 (미국 특허 2건 등록 완료)

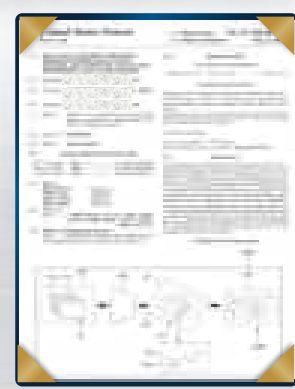
종류	등록 번호	내용	국가
특허	10-2765662	데이터 클리닉 방법, 데이터 클리닉 방법이 저장된 컴퓨터 프로그램 및 데이터 클리닉 방법을 수행하는 컴퓨팅 장치	KR
특허	10-2535687	이미지에 대응되는 형상 이미지를 제공하기 위한 이미지 처리 방법 및 그러한 방법을 수행하는 전자 장치	KR
특허	10-2535688	픽셀 기반의 이미지 처리 방법 및 그러한 방법이 구현된 사용자 인터페이스를 포함하는 전자 장치	KR
특허	10-2535692	복수의 이미지에 포함되는 픽셀 사이의 대응 관계를 반영하는 이미지를 획득하기 위한 이미지 처리 방법 및 그러한 방법을 수행하는 전자 장치	KR
특허	10-2519094	합성 데이터 생성을 위한 가상 환경을 제공하는 전자 장치, 전자 장치의 동작 방법 및 전자 장치를 포함하는 시스템	KR
특허	10-2719240	가상 데이터를 생성하기 위한 방법 및 그러한 방법이 구현된 컴퓨팅 장치	KR
특허	10-2739394	가상 데이터를 이용한 평가 방법 및 그러한 방법이 구현된 컴퓨팅 장치	KR
특허	10-2719246	가상 데이터를 이용한 인공지능 모델의 학습 방법 및 그러한 방법이 구현된 컴퓨팅 장치	KR
특허	10-2665956	가상 데이터를 처리하기 위한 유저 인터페이스 제공 방법 및 그러한 방법이 구현된 컴퓨팅 장치	KR
특허	11,748,447	DATA CLINIC METHOD, COMPUTER PROGRAM STORING DATA CLINIC METHOD, AND COMPUTING DEVICE FOR PERFORMING DATA CLINIC METHOD	US
특허	11,868,435	DATA DIAGNOSIS METHOD, A COMPUTER PROGRAM STORING THE DATA DIAGNOSIS METHOD, AND COMPUTING DEVICE FOR PERFORMING DATA DIAGNOSIS METHOD	US



국내특허등록증



미국특허등록증



페블러스는 국내 산업 전반의 다양한 프로젝트 수행 경험을 확보하였으며, 데이터 품질 관리 및 학습데이터 구축 수행 실적으로 보유하고 있는 전문업체입니다.

현대자동차

차량 운행중 센터 데이터를 통한 위치 추정 AI 개발

문제점

- 1. 낮은 성능
- 2. 공간 제약에 따른 데이터 취득의 어려움

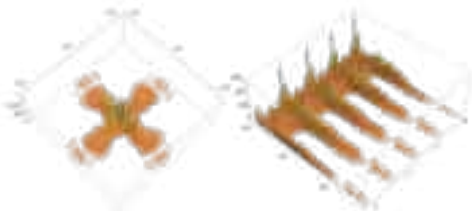
솔루션

- 1. 데이터 품질 진단
- 2. 정밀 타기팅 합성데이터 생성
- 3. AI 참조모델 개발

솔루션

- 1) 데이터 취득 과정을 **15일에서 1시간 이하로 단축**
- 2) 공간섭외 및 인력투입에 필요한 **비용 절감**
- 3) **200% 이상 AI 성능 향상**

규모 2022년 제로원 프로젝트. 5개월, 5천만원



한화비전

객체검출 및 상황인지를 위한 AI 감시 카메라 개발

문제점

- 1. 고객향 데이터 품질 입증 자료 부재
- 2. 라벨링 부재

솔루션

- 1. 최적 신경망 선별을 통한 오토라벨링 제공
- 2. 데이터 클리닉을 통한 품질 진단

솔루션

- 1) 고품질 자동 라벨링을 통한 시간 및 비용 절감
- 2) 데이터셋의 품질 상태 확인으로 **차세대 데이터 거버넌스/파이프라인 필요성 인지**

규모 2024년 PoC 프로젝트. 5개월, 5천만원



자이언트스텝

가상 아이돌 제작을 위한 라이브 페이스 캡처 시스템 개발

문제점

- 1. AI 학습 시간 병목
- 2. 중복이 심한 200만 건의 합성데이터

솔루션

- 1. 데이터 클리닉을 통한 품질 진단
- 2. 합성데이터 생성
- 3. 인공지능 참조모델 개발

솔루션

- 1) 15% 데이터 경량화로 **학습시간 1주일 → 1일로 단축**
- 2) S사 가상 아이돌 제작에 활용
- 3) SIGGRAPH Asia 2025 학회 발표

규모 2024년 데이터 클리닉 적용. 12개월, 1억원



해병대

해양 경계감시 AI 개발

문제점

해무, 파도 등 다양한 환경 및 북한군 객체 데이터 취득 어려움

솔루션

컴퓨터 그래픽스, 생성모델을 활용한 고품질 합성데이터 생성

솔루션

- 1) 실제 취득이 어려운 13.5만건 합성이미지 생성
- 2) **비주일 튜링 테스트 만점 수준** (49점으로 통과. 실제와 합성을 구별하지 못할 때 50점)

규모 2024년 과기부 프로젝트, 6개월, 2.5억원



페블러스는 국내 산업 전반의 다양한 프로젝트 수행 경험을 확보하였으며, 데이터 품질 관리 및 학습데이터 구축 수행 실적으로 보유하고 있는 전문업체입니다.

연구개발과제명	연구개발기간	수행내용	중앙행정기관 (전문(전담)기관)
환경 지속가능성을 위한 재활용 분류 인공지능 학습데이터 구축	22.06.13~22.11.30	자원 재활용 인공지능 합성데이터 생성	대전창조경제혁신센터
비대면 스포츠 앱 데이터 분석 서비스	22.05.09~22.12.31	비대면 스포츠 앱 사용자 데이터 분석	문화체육관광부
페플라ستيك 재활용 분류를 위한 인공지능 학습데이터 구축 사업화	22.05.01~23.02.28	인공지능 합성데이터 구축 사업화	창업진흥원
정밀 타게팅 가상데이터 생성솔루션	21.11.01~23.12.31	3세대 인공지능 합성데이터 솔루션 구축	중소벤처기업부 (중소기업기술정보진흥원)
삼차원 얼굴 데이터 취득을 위한 뉴럴페이스 모델링 및 인터페이스 개발	23.09.01~24.08.31	디지털 휴먼 표정 생성을 위한 얼굴 데이터 구축 및 임베딩 기술 사업화	산업통상자원부 (한국산업기술진흥원)
정보통신·방송 기술개발사업 (예외 상황 합성데이터 생성 및 인공지능 예측 모델 고도화 기술개발)	24.04.01~27.12.31	데이터가 극히 부족하거나 전혀 없는 예외상황에서도 인공지능 예측 모델에 활용할 수 있는 “시뮬레이션 기반 합성데이터 생성 기술 및 생성된 데이터를 활용한 예측모델 고도화 기술” 개발	과학기술정보통신부 (정보통신기획평가원)
군 경계 작전 환경 데이터 (초거대시 확산 생태계 조성사업)	24.07.01~24.12.31	AI 학습용 이미지 합성데이터 생성	한국지능정보사회진흥원 (전담기관)

AI 학습데이터 품질에 관한 최고 전문가들을 투입하여 사업을 성공적으로 수행하겠습니다.



사업총괄책임자(PM)

이정원 부대표, 프로젝트 총괄 책임

자문그룹

데이터 거래사(사내 1인)

데이터인터랙션팀 2명

사전 컨설팅

솔루션개발팀 2명

데이터 품질 진단

품질 개선

데이터사이언스팀 2명

데이터 품질 진단

품질 개선

합성데이터팀 1명

품질 개선

이름	그룹	담당업무	해당분야 경험
이정원	사업총괄책임자	사업 총괄	데이터 분야 사업 PM
이주행	자문그룹	데이터 컨설팅 및 자문	데이터 거래사
김중현	데이터인터랙션팀	사전 컨설팅	데이터 시각화 도구 개발
천수비	데이터인터랙션팀	사전 컨설팅	데이터 시각화 도구 개발
정지엽	솔루션개발팀	데이터 품질 진단/개선	데이터 클리닉 개발
허희석	데이터사이언스팀	데이터 품질 진단/개선	품질 진단 기술 개발
김승현	데이터사이언스팀	데이터 품질 진단/개선	품질 진단 기술 개발
장홍탁	합성데이터팀	데이터 품질 개선/ 합성데이터 생성	합성데이터 생성 기술 개발

내 용	7월			8월			9월			10월			11월		
수요기업 모집	수요기업 모집														
사전 컨설팅			사전 컨설팅												
1차 품질진단/개선 (5개사)				컨설팅	품질진단		품질개선								
2차 품질진단/개선 (5개사)					컨설팅	품질진단									
3차 품질진단/개선 (5개사)							컨설팅	품질진단							
4차 품질진단/개선 (5개사)									컨설팅	품질진단	품질개선				
후속지원															후속지원

데이터 품질진단 및 개선

AI 대전환 시대의 핵심 과제, 대구 지역 기업의 데이터 품질 고민 저희가 함께 해결하겠습니다.

고품질 AI-Ready Data로 산업 혁신을 이끄는 최적 파트너

 대구디지털혁신진흥원
Daegu Digital Innovation Promotion Agency

