

인천일보

시사 > 전체기사

인천 스타트업파크 12월 중순 오픈 빅데이터 도입

플랫폼과 프로그램 도입, 인천 스타트업 파크 입주기업 새로운 비즈니스 모델 개척 효과 기대



소셜 빅데이터 분석 프로그램을 통한 인천스타트업파크 이미지, 인천경제자유구역청 제공

인천경제자유구역청은 인천 스타트업파크 입주 기업들에게 빅데이터를 자유롭게 활용하고 인공지능 개발이 쉬운 환경을 제공하기 위해 이 같은 플랫폼과 프로그램을 도입 중에 있다고 26일 밝혔다.

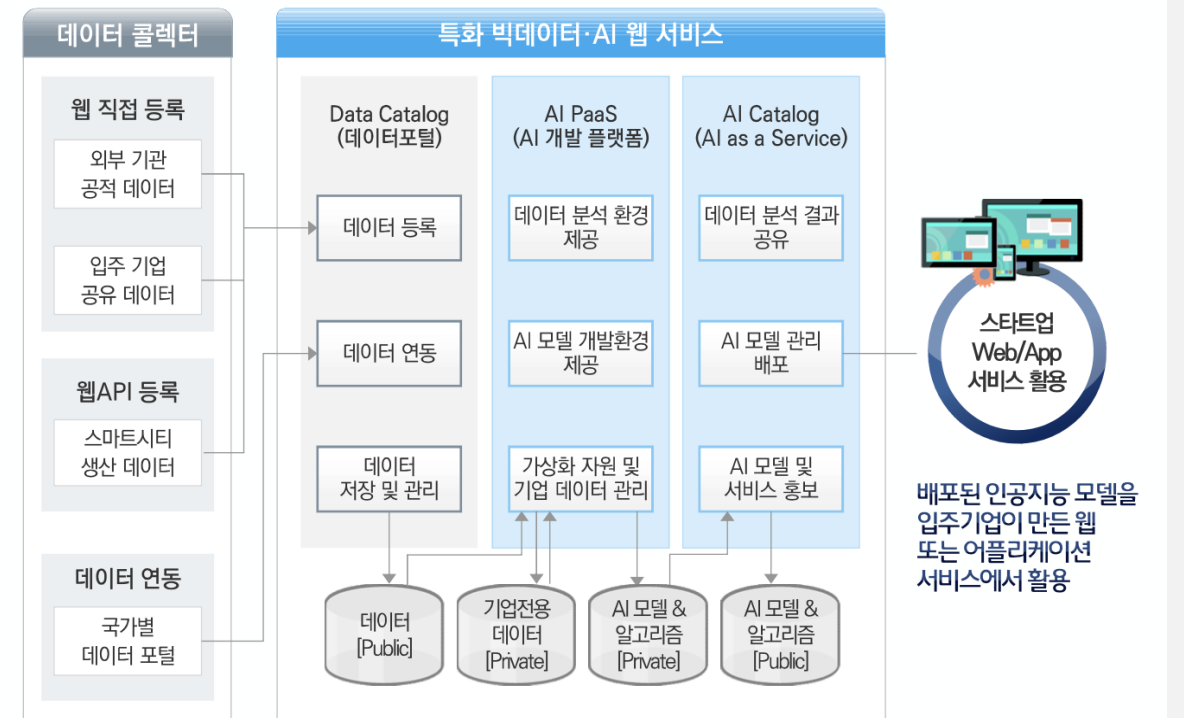
빅데이터와 인공지능 관련 기업은 방대한 데이터를 확보하고 각각의 서버를 통해 이를 다루는데 막대한 비용과 인력이 소요된다. 빅데이터-인공지능 플랫폼은 서비스를 개발하는데 드는 비용을 절감하고 효율적으로 데이터를 다룰 수 있도록 해 입주 기업에게 큰 도움이 될 것으로 인천경제청은 기대하고 있다.

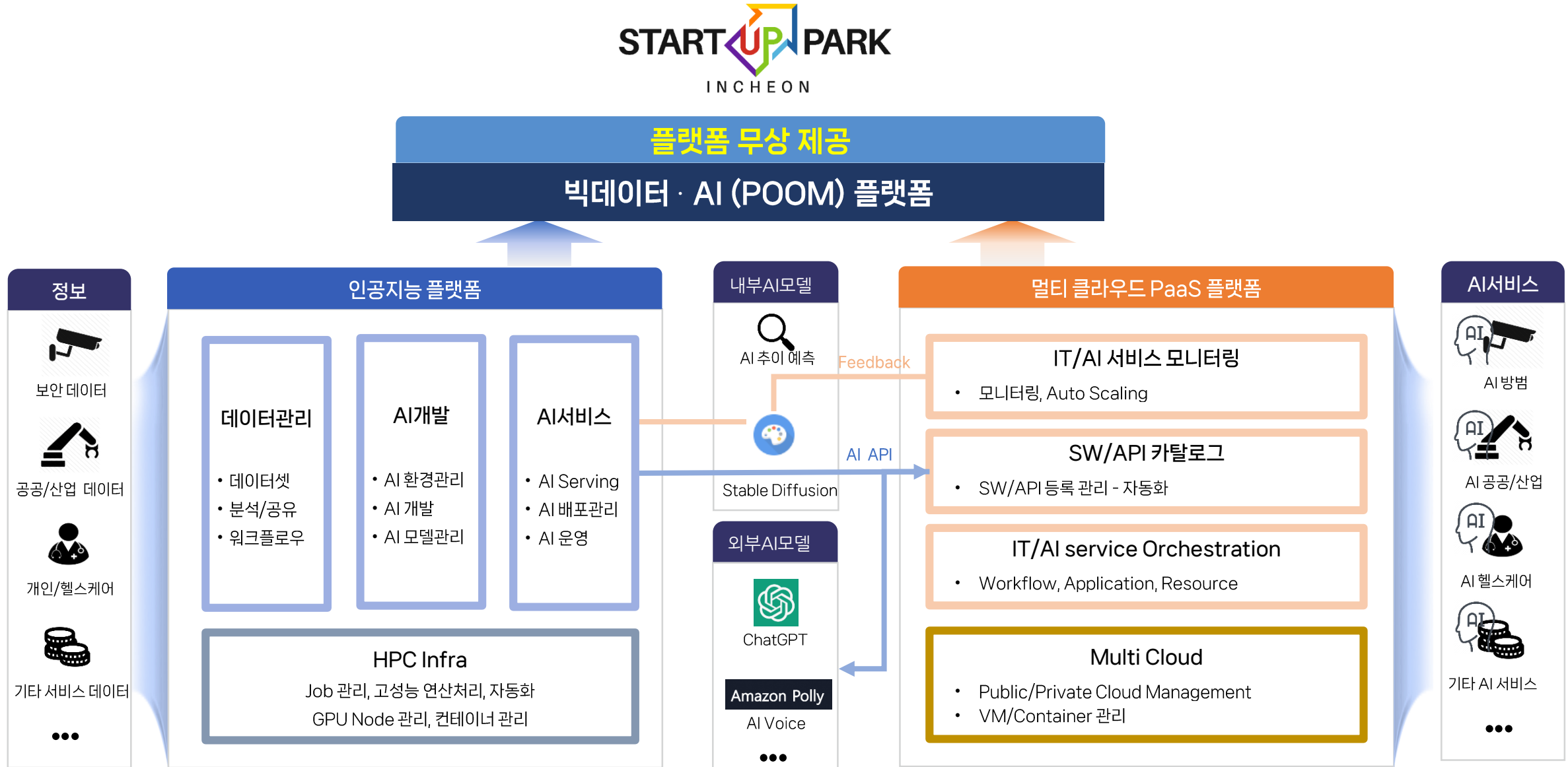
빅데이터-인공지능 플랫폼은 데이터의 생산부터 수집, 관리, 공유를 손쉽게 하고 인공지능 개발자가 데이터 학습과 알고리즘 개발에만 집중할 수 있는 인프라다.

STARTUP PARK INCHEON

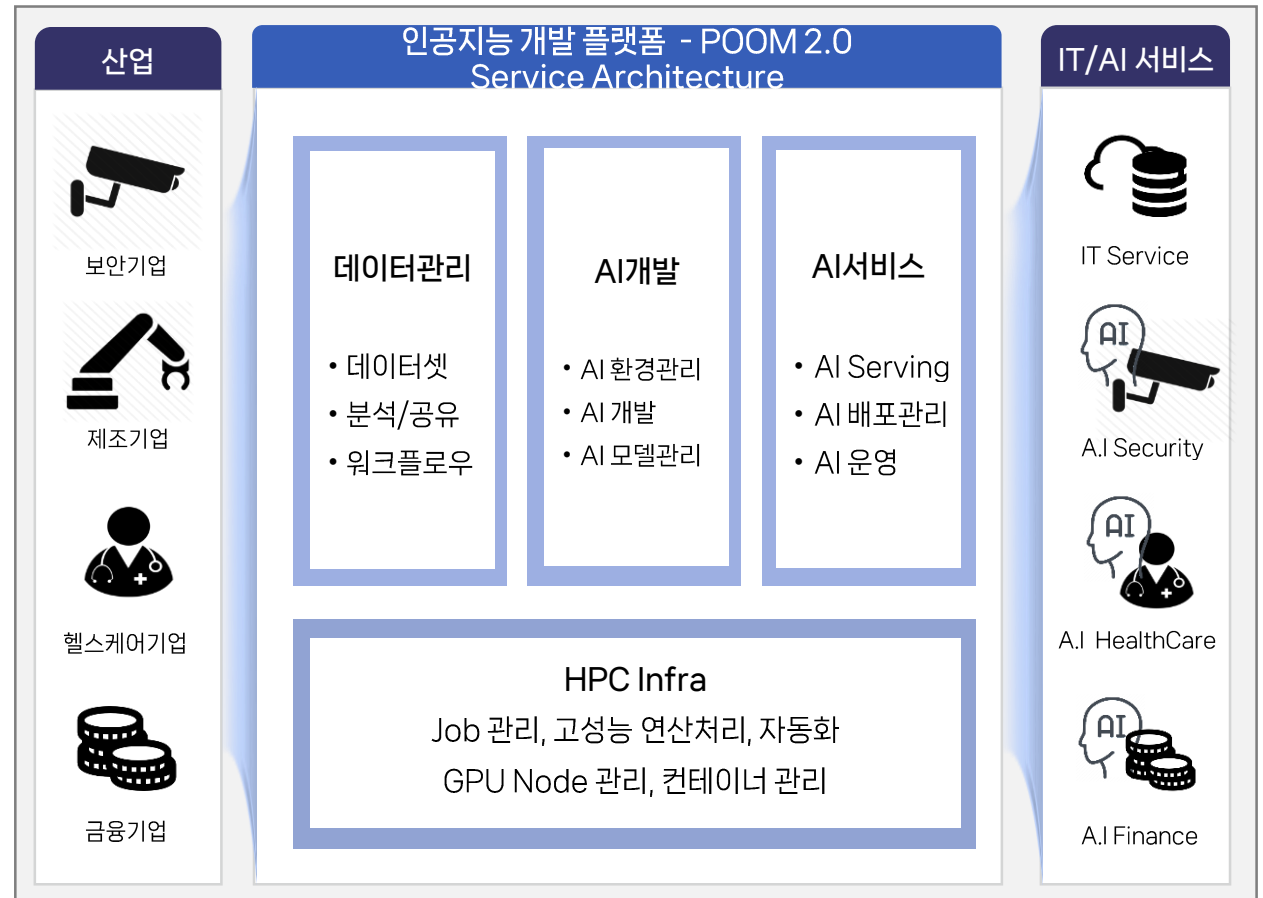
인천스타트업파크는 스타트업이 데이터를 활용하여 신규 비즈니스의 창출, 서비스 개발을 진행할 수 있도록, 데이터 수집에서부터 인공지능 개발, 실증까지 AI개발/서비스가 순차적으로 연결될 수 있는 제공하는 통합 AI플랫폼을 운영 (2021년 부터 2025년 까지 누적 지원기업 73개사)

[인천 스타트업파크 빅데이터/AI 서비스 구성도]





인천스타트업파크 "AI 포탈 서비스" 는 다수의 사용자가 효율적인 인프라 환경에서
데이터 관리부터 AI모델 개발, 배포할 수 있는 인공지능 개발 플랫폼(MLOps)입니다.



인천스타트업파크 "PaaS 포탈 서비스" 는 다수의 사용자가 효율적인 인프라 환경에서 클라우드 기반의 다양한 IT AI 서비스를 위한 최적의 (DevOps) 환경을 제공합니다.

표준화된 클라우드 자원 관리



- 각종 클라우드 자원에 대한 인터페이스를 표준화 UI 로 제공
- 온프레미스 인프라 통합 제어 관리
- 모니터링, 스케줄링, 빌링리포트 등 리소스 효율성 개선

Low Code 서비스 빌딩



- GUI 기반의 워크플로우와 Web IDE로 Low Code 개발 환경 제공
- SW, API 등 카탈로그 정의를 통해 높은 재사용성 및 개발기간 단축
- 별도의 서버 환경 구축이 필요 없는 Serverless 아키텍처 지원

서비스 자동화 오케스트레이션



- 워크플로우 기반 카탈로그를 제공하여 서비스 실행과 배포 자동화
- 마이크로서비스 아키텍처 기반으로 CI/CD 및 고가용성 지원
- 각 기능요소 단위 API를 제공하여 서비스 확장이 용이

서비스 개선을 위한 부가서비스



- 서비스 프로젝트 단위 실시간 리소스 모니터링 및 자원 관리
- 멀티 테넌시 구조로 사용자별 권한 관리와 협업 환경 제공
- 리소스 상황에 대한 실시간 Notification 및 이력 관리

폼 플랫폼 주요기능 (1/2)

내용

· AI 서비스 개발을 위한 개발 프레임워크 제공 (프레임워크 : Python, Jupyter, PyTorch, Keras, TensorFlow 등)

① 데이터 구축/처리, ② AI모델 개발, ③ AI모델 배포/운영, ④ HPC 인프라 관리 등 각 **AI 개발 단계별 최적의 환경**을 제공

POOM
플랫폼
주요 기능
(AI Portal)



폼 플랫폼 주요기능 (2/2)

내용

- 클라우드 기반의 다양한 IT · AI 서비스를 위한 최적의 DevOps 환경을 제공

PaaS Portal Workflow 기능을 통해 기업만의 AI/IT 서비스를 만들 수 있습니다.

POOM
플랫폼
주요 기능
(PaaS Portal)

어플리케이션 개발

자유로운 어플리케이션 라이브러리

- 실제 서비스 운영에 필요한 다양한 형태의 어플리케이션을 등록
- 등록한 어플리케이션을 빠르게 재사용 가능
- 지속적인 수정 및 버전관리

리소스 카탈로그 생성 기능

- 필요한 자원을 즉시 사용할 수 있도록 자원이 규격화되어 제공
- 리소스 카탈로그상에서 서비스 즉시 실행
- 서비스 카탈로그에서 실행 기능 지원

서비스 배포 및 운영

서비스 중단 없는 일관된 배포 환경

- 어플리케이션 코드를 도커 이미지로 Export 하여 마들웨어와 함께 즉시 실행할 수 있는 무중단 서비스 배포 기능을 제공
- 외부 데이터베이스나 직접 구성한 데이터베이스를 데이터 스토어에 등록/사용 가능

GUI 방식의 워크플로우 설계

- 원하는 어플리케이션을 조합하여 Flow Chart를 생성가능
- 워크플로우 디자인 캔버스 제공
- 서비스 생성 관리 기능 제공

서비스 통합관리

인스턴스 및 자원 관리

- 스케줄 기능을 통한 서비스 자동화와 자원 사용량에 따른 Auto Scaling 기능을 지원
- 단일 Viewpoint에서 여러 자원 관리 기능
- 지속적인 수정 및 버전관리

Workflow 제공

- GPT-3 Sample Workflow 제공(참여기업 재사용 가능)
- 참고용 최종 개발 예시 제공(PoomGPT)
- 기업이 직접 수행할 수 있게 하는 실무형 Workflow 교육 제공

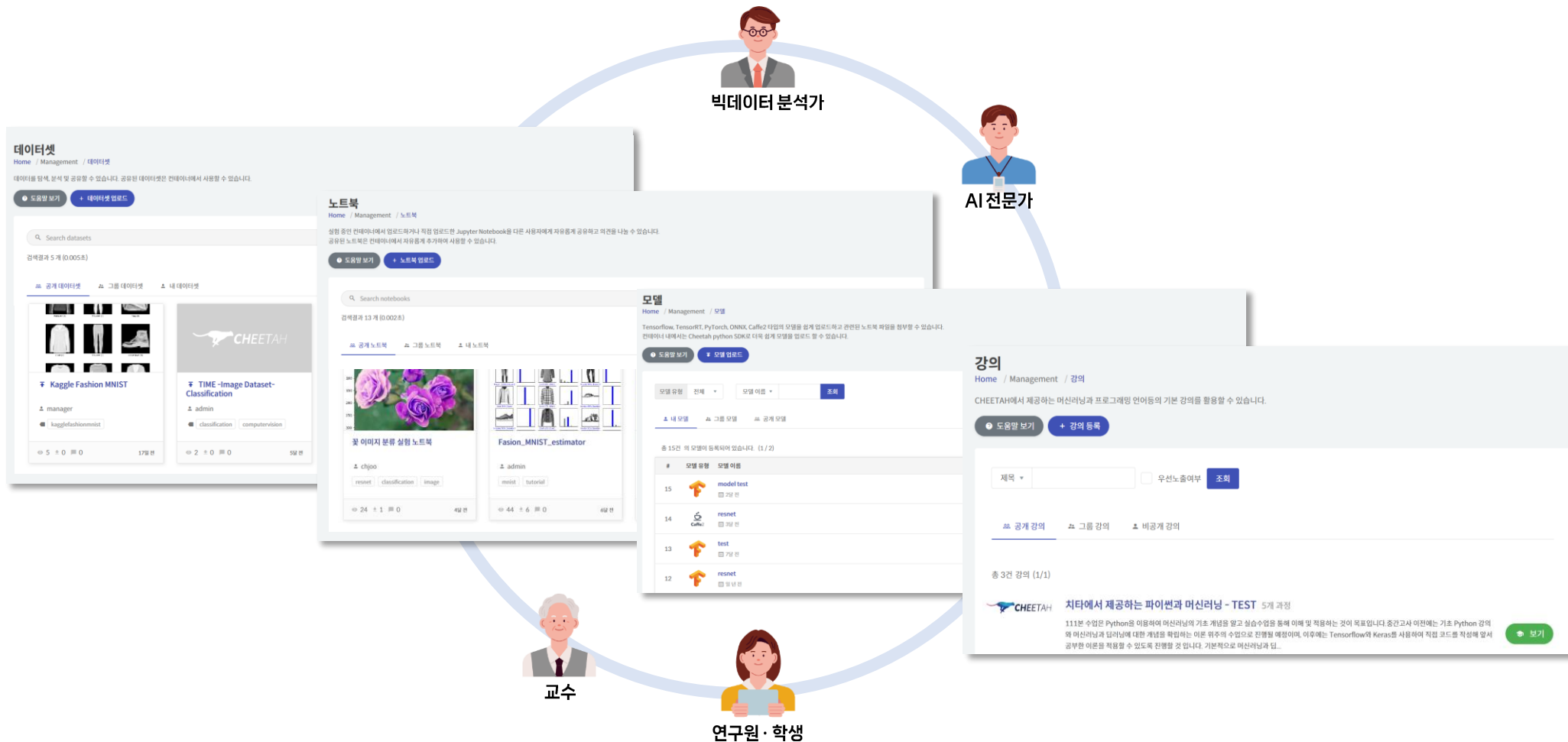
인공지능 개발에 필요한 최신 라이브러리와 프레임워크 환경을 **클릭만으로 1분 안에 생성** 합니다.

워크스페이스 생성에 필요한 **다양한 라이브러리와 프레임워크가 버전별로 탑재**되어 자유롭게 선택할 수 있습니다.



	Ubuntu	Python	JupyterLab	VSCode	XFCE	CUDA	cuDNN	TensorFlow	Keras	PyTorch	torch	horovod	Mxnet	R	RStudio	Java	TinyTex	Tidyverse	MATLAB
20.04	3.8	3.0.16	3.11.1	4.12.0	11.3	8	2.7.0	2.7.0	3.8										
20.04	3.8	3.0.16	3.11.1	4.12.0	11.3	8	2.6.0	2.6.0	3.8										
20.04	3.8	3.0.16	3.11.1	4.12.0	11.3	8	2.5.0	2.4.3	3.8										
20.04	3.8	3.0.16	3.11.1	4.12.0	11.3	8			3.8	1.12.1									
20.04	3.8	3.0.16	3.11.1	4.12.0	11.3	8	2.7.0		3.8	1.12.1		0.25							
20.04	3.8	3.0.16	3.11.1	4.12.0	11.1	8	2.4.0	2.4.3	3.8										
20.04	3.8	3.0.16	3.11.1	4.12.0	11.1	8			3.8										
20.04	3.8	3.0.16	3.11.1	4.12.0	11.1	8			3.8										
20.04	3.8	3.0.16	3.11.1	4.12.0	11.1	8	2.4.0	2.4.3	3.8										
20.04	3.8	3.0.16	3.11.1	4.12.0	11.1	8			3.8	1.8.0									
18.04	3.7	2.2.8	3.6.2	4.12.0	10.1	7	2.3.0	2.4.0	3.7										
18.04	3.7	2.2.8	3.6.2	4.12.0	10.1	7	2.2.0	2.4.0	3.7										
18.04	3.7	2.2.8	3.6.2	4.12.0	10.1	7	2.1.0	2.3.1	3.7										
18.04	3.7	2.2.8	3.6.2	4.12.0	10.1	7			3.7	1.7.0									
18.04	3.7	2.2.8	3.6.2	4.12.0	10.1	7			3.7	1.6.0									
18.04	3.7	2.2.8	3.6.2	4.12.0	10.1	7			3.7	1.5.1									
18.04	3.7	2.2.8	3.6.2	4.12.0	10.1	7			3.7	1.4.0									
18.04	3.7	2.2.8	3.6.2	4.12.0	10.1	7	2.3.0		3.7	1.6.0	0.22	1.6.0							
18.04	3.7	2.2.8	3.6.2	4.12.0	10.1	7	2.1.0		3.7	1.4.0	0.19.1	1.6.0							
18.04	3.6	2.2.8	3.6.2	4.12.0	10.0	7	1.15.0	2.3.1	3.7										
18.04	3.6	2.2.8	3.6.2	4.12.0	10.0	7	1.14.0	2.2.5	3.7										
18.04	3.6	2.2.8	3.6.2	4.12.0	10.0	7	2.0.0	2.3.1	3.7										
18.04	3.6	2.2.8	3.6.2	4.12.0	10.0	7			3.7	1.4.0									
18.04	3.6	2.2.8	3.6.2	4.12.0	10.0	7			3.7	1.3.0									
18.04	3.6	2.2.8	3.6.2	4.12.0	10.0	7			3.7	1.2.0									
18.04	3.6	2.2.8	3.6.2	4.12.0	10.0	7	2.0.0		3.7	1.2.0	0.19.0	1.5.0							
18.04	3.6	2.2.8	3.6.2	4.12.0	10.0	7	1.14.0		3.7	1.2.0	0.18.1	1.5.0							
16.04	3.5	2.2.8	3.6.2	4.12.0	9.0	7	1.12.0	2.2.4	3.5										
16.04	3.5	2.2.8	3.6.2	4.12.0	9.0	7			3.5	1.1.0									
16.04	3.5	2.2.8	3.6.2	4.12.0	9.0	7	1.12.0		3.5	1.1.0	0.16.4	1.4.1							
20.4	3.8		4.7.0	4.12.0	11.1					4.2.1	2022.07.2+	1.8	venv	venv					
20.4	3.8		4.7.0	4.12.0	11.1					4.2.1	2022.07.2+	1.8	venv	venv					License Required

Data Set, Jupyter Notebook, AI 모델 등의 AI Asset을 다른 사용자와 자유롭게 공유하고 협업 할 수 있으며,
사내 AI비전문가분들도 잘 활용할 수 있도록 강의 기능을 활용하여 AI 교육은 물론 실습까지 가능합니다.



ML 파이프라인 자동화를 제공하여 모델을 지속적으로 학습시킬 수 있고
개발된 AI 모델은 몇번의 클릭만으로 **손쉽게 배포, 활용** 할 수 있습니다.



빠른 AI 개발 환경(워크스페이스) 구축

CUDA, TensorFlow, Keras Version 등의 복잡한 설치과정이 없이 몇 번의 클릭으로 원하는 인공지능 개발 워크스페이스 환경 구축이 가능합니다.

1 GPU/CPU 자원 선택

컨테이너
Home / Management / 컨테이너
☰ 자원

* 자원

✓ GPU

ON_PREMISE, GPU

○ NVIDIA-GeForce-RTX-2080-Ti

NVIDIA, turing, @ 11.019Mi

VCPU 8.0 core, Memory 62.0 Gi

• CPU

ON_PREMISE, GPU

○ NVIDIA-A100-PCIE-40GB

NVIDIA, ampere, @ 40.536Mi

VCPU 30.0 core, Memory 56.0 Gi

② 내 GPU 사용 현황

② 노드의 할당 가능한 GPU 현황

■ 사용중 ■ 사용 중인 GPU에 할당 ■ 할당됨 ■ 대기중

0 개

1 개

2 이미지 버전 선택

☞ 이미지

* 이미지 선택

✓ 빌트인 이미지

• 사용자 이미지

cheetah/gpu-workspace:ubuntu20.04-cuda11.3-cudnn8

cheetah/gpu-workspace:ubuntu20.04-cuda11.3-cudnn8

cheetah/gpu-workspace:ubuntu20.04-cuda11.1-cudnn8

cheetah/gpu-workspace:ubuntu18.04-cuda10.1-cudnn7

cheetah/gpu-workspace:ubuntu18.04-cuda10.0-cudnn7

cheetah/gpu-workspace:ubuntu16.04-cuda9.0-cudnn7

cheetah/gpu-workspace:r-ml-verse-4.2.1-cuda11.1

3 각 이미지별 최신 라이브러리와 분석 툴 기본 설치 확인

가상 환경(커널) 옵션

* 가상 환경

tf2.3.0-keras2.4.0-py3.7-cuda10.1

tf2.3.0-keras2.4.0-py3.7-cuda10.1

tf2.2.0-keras2.4.0-py3.7-cuda10.1

tf2.1.0-keras2.3.1-py3.7-cuda10.1

torch1.7.0-py3.7-cuda10.1

torch1.6.0-py3.7-cuda10.1

torch1.5.1-py3.7-cuda10.1

torch1.4.0-py3.7-cuda10.1

horovod0.22.0-tf2.3.0-torch1.6.0-mxnet1.6.0-py3.7-cuda10.1

horovod0.19.1-tf2.1.0-torch1.4.0-mxnet1.4.0-py3.7-cuda10.1

컨테이너 환경

구성 환경

Ubuntu 20.04

Python 3.8

JupyterLab 3.0.16

VSCode 3.11.1

XFCE 4.12.0

CUDA 11.3

cuDNN 8

TensorFlow venv

4 볼륨/데이터셋 연결 후 워크스페이스 생성 완료!

볼륨

볼륨 연결 추가

구분 볼륨(볼륨명/용량/마운트 패스)

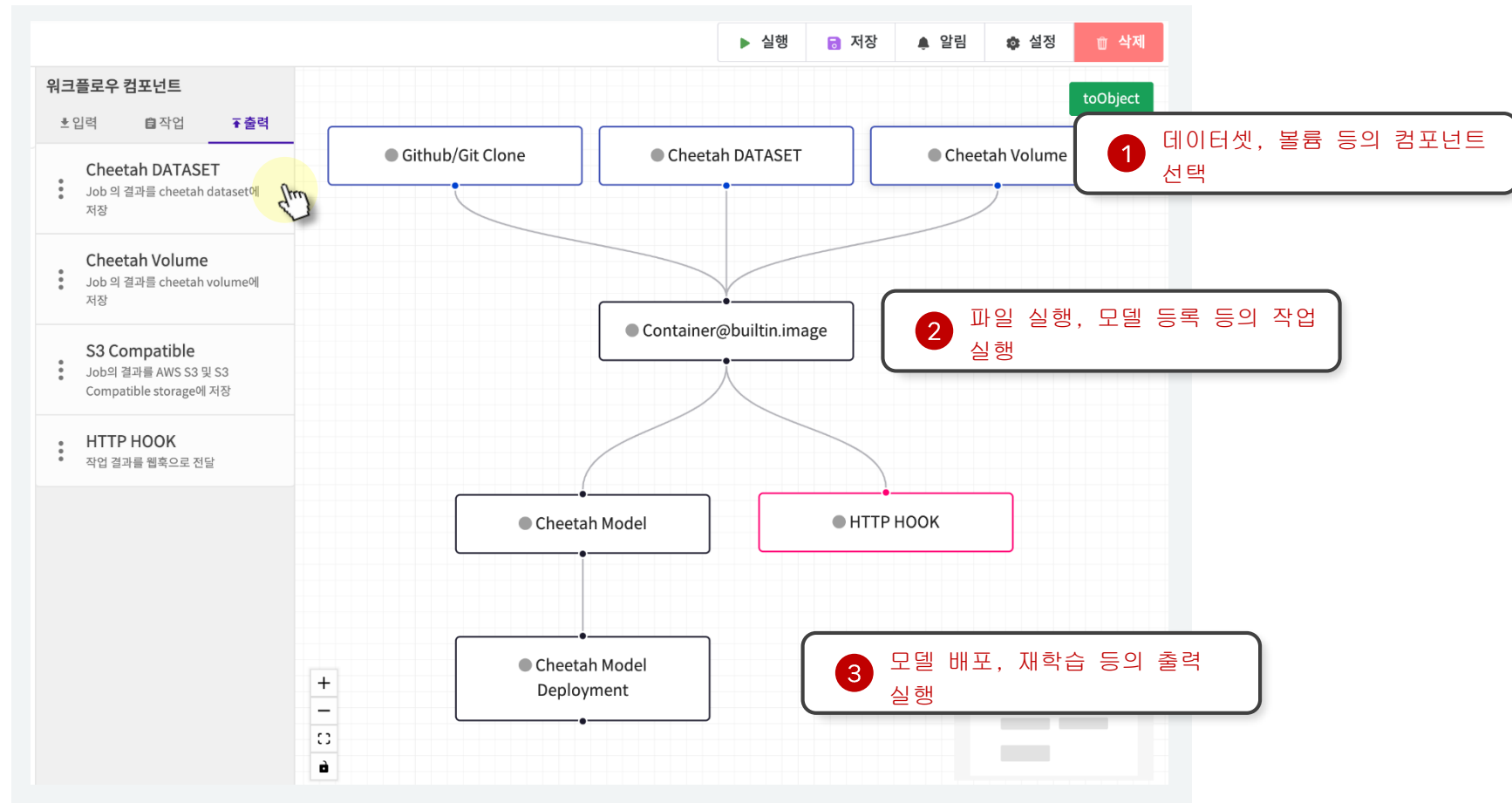
기본 Home

10 GiB / \$USER_HOME

CEPH 수주 500500 삭제

ML 파이프라인 자동화

ML 파이프라인 자동화를 통해 데이터 처리부터 모델 배포까지, **반복적이고 소모적인 작업을**
더 쉽고 효율적으로 **처리하며 빠르게 모델을 배포**할 수 있습니다.



Job Scheduling으로 효율적인 자원 사용

멀티 노드 클러스터링 및 다양한 프레임워크를 통해 Job을 여러 노드에서 병렬로 실행 가능하여
AI-ML 연산을 초고속으로 처리할 수 있습니다.

1 분산 트레이닝 Type 선택

2 프레임워크 선택

3 실행 일자/시각 설정

4 사전 검증

test job
수수 (soosoo), 2023-02-16 10:39

컨테이너 정보

자원	NVIDIA-GeForce-RTX-2080-Ti x 1
Distribution Type	Distributed GPU
Image	cheetah/gpu-workspace:ubuntu20.04-cuda11.3-cudnn8
Kernel	tf2.7.0-keras2.7.0-py3.8-cuda11.3
볼륨	/home/jovyan 10GB /home/jovyan/job_sc 2GB

잡스케줄링 실행 정보

실행 타입	PYTHON3
실행 경로	/home/jovyan/job_sc
실행 명령어	python3 Fasion_MNIST_estimator_job.py
예정된 런타임	2023-02-24 01:30 ~ 잡스케줄링 종료시간이 정의되지 않음
런타임	확정된 잡이 아직 수행되지 않음

잡스케줄링 이력

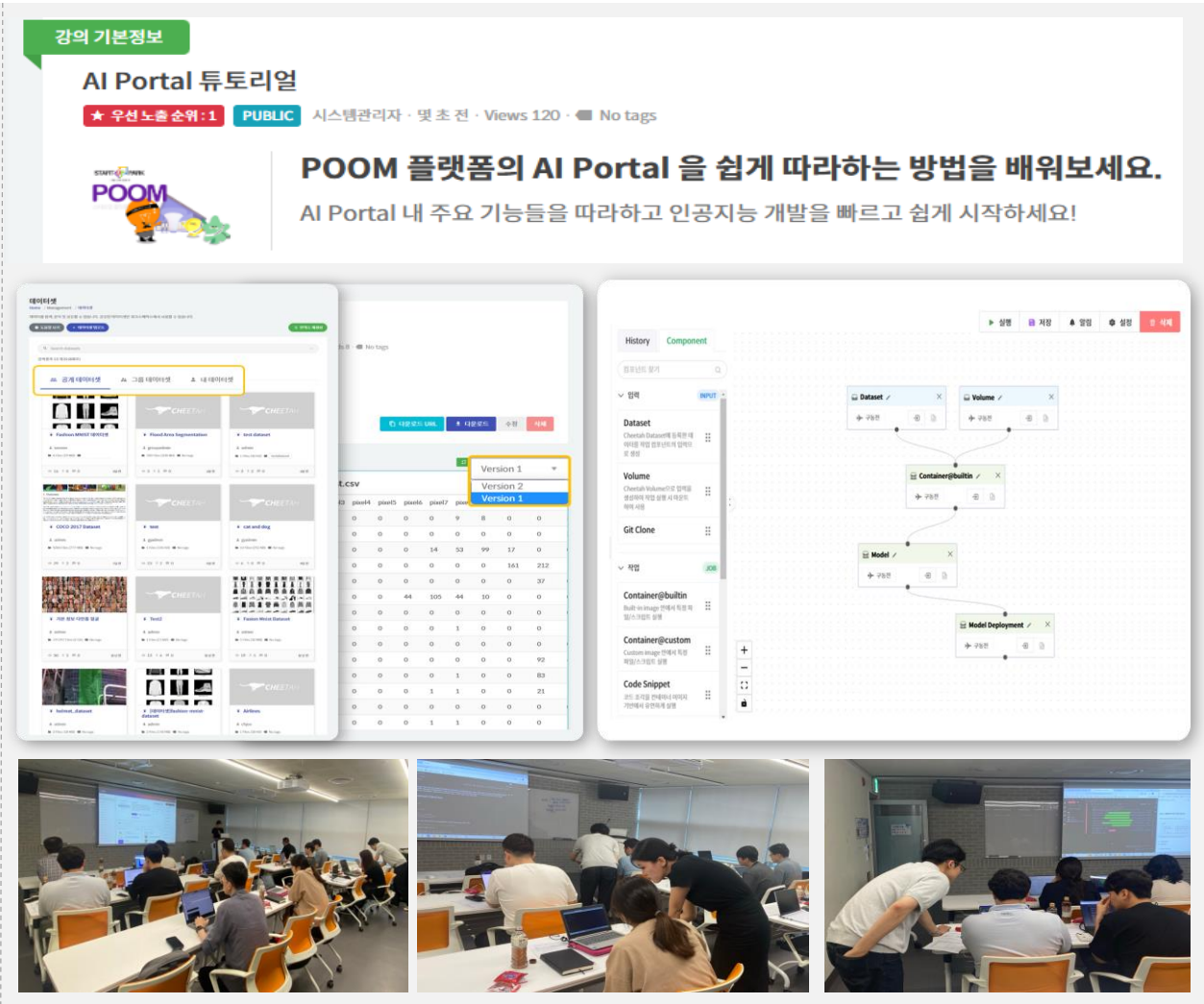
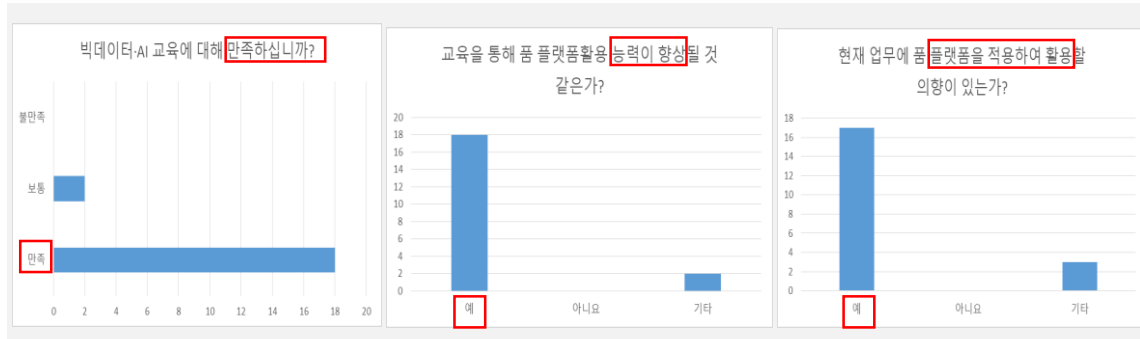
파드명	단계	상태	시작일	종료일	로그	상세
job-736f6f736f6f-3fmcag-jbb2s	테스트	종료(성공)	2023-02-16 10:40:22	2023-02-16 10:41:34		

테스트

[교육 내용]

- 장소 : 인스타 1 4층 교육실
- 교육목적 :
 - 빅데이터 AI 기술활용을 위한 플랫폼 실습 교육
- 교육 내용 (NVIDIA-DLI (Deep Learning Institute)
 - AI개발 실습 (컨테이너 생성, 볼륨, Jupyter notebook 활용 실습)
 - AI Portal 을 통한 모델 훈련, 가상환경 생성, 모델 배포 실습
 - Hugging Face 에서 LLM 모델 생성 및 활용을 위한 실습
 - Application 개발 및 실행, 머신러닝 파이프라인 자동화 구현 실습
 - AppDev 를 통한 코드 실습 (Work-Flow 구성,실행,서빙모델 연동)

[교육 설문 조사]



스타트업 초기에는 데이터를 구하기가 쉽지 않습니다.
초기에는 회사 제품/서비스에 사용할 수 있는 데이터를 확보하여, AI개발을 시작해 나가야 합니다.

Startup	데이터 구축	데이터 처리	인공지능 개발	인공지능 서비스
Have DATA	<ul style="list-style-type: none"> 공개 데이터 수집, 판매 데이터 구매 	<ul style="list-style-type: none"> 오픈소스 기반 데이터 전처리 	<ul style="list-style-type: none"> 오픈소스 AI모델 파인튜닝 	<ul style="list-style-type: none"> AI를 제품/서비스 탑재(Embedded) AI 서비스(Web)
No DATA	<ul style="list-style-type: none"> 특화 데이터 수집 <ul style="list-style-type: none"> Web/APP, 외부기관협력 	<ul style="list-style-type: none"> 특화 데이터 복수 단계 전처리 	<ul style="list-style-type: none"> 오픈소스 AI모델 파인튜닝 반복/강화 AI모델 고도화 	<ul style="list-style-type: none"> AI를 제품/서비스 탑재(Embedded) AI 서비스(Web)

Data Catalog Portal

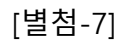
✓ 등록된 데이터에서 자신에게 필요한 데이터를 획득

AI Portal

- ✓ 데이터를 처리하고
- ✓ 인공지능 개발 (고성능 GPU제공)

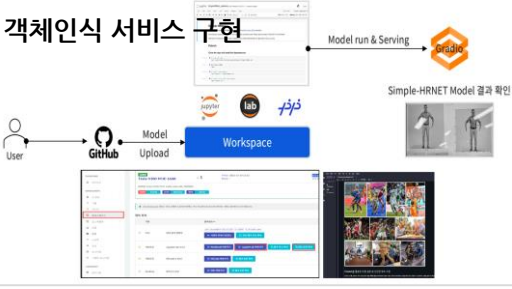
PaaS Portal

✓ 클라우드에 서비스를 배포/운영



POOM 플랫폼을 활용한 사례

순번	기업명	발표자	발표 요약 [플랫폼 활용내용]
1	서울의료정보연구소 	성**	디지털헬스 데이터 생성·추출·Framework Use Case [AI Portal / GPU 2개 활용]
2	그린플래그 	안**	폼 플랫폼과 ChatGPT를 활용한 그림책 제작 Use Case [PaaS Portal / App/Dev 활용]
3	비바이노베이션 	김**	건강데이터 추적관찰 GPT 활용 Use Case [AI Portal / GPU 3개 활용]
4	도터 	김**	OCT AI 분석 & Train Model 학습 Use Case [AI Portal / GPU 1개 활용]
5	더좋은운동으로 	고**	Human Pose Estimation Use Case [PaaS Portal / App/Dev 활용]
6	몬드리안에이아이 	구**	GPU 중개 플랫폼 개발 Use Case [AI · PaaS Portal / GPU 1개 활용]



스타트업의 AI 기술발전을 위해 기업 컨설팅 및 AI 파일럿 개발을 지원 하였습니다.

개발 시 완료된 모델은 플랫폼내에서 API 를 통해 테스트 가능 합니다.



구분	상세							
AI Portal	운영체제	리눅스 (Ubuntu 22.04, 20.04) 제공						
	개발환경	프레임워크	개발 언어	개발 환경	데스크탑 환경	병렬처리도구	CUDA	cuDNN
		Tensorflow - 2.9.0, 2.7.0, 2.6.0, 2.5.0, 2.4.0 , 2.16.1, 2.15.0, 2.14.0	Python - 3.11, 3.10, 3.9, 3.8 JAVA - 1.8	JupyterLab - 4.1.4, VsCode - 4.22	XFCE - 4.12.0	OpenMPI - 4.1.2 Horovod - 0.25.0, 0.28.1	11.1, 11.3, 11.5, 11.6, 11.7, 11.8, 12.0, 12.1	7, 8
		Pytorch - 1.11.0, 1.12.1, 1.13.1, 1.8.0, 1.9.0, 2.0.1, 2.1.2, 2.2.2, 2.3.0						
		Keras - 3.3.3, 2.9.0, 2.7.0, 2.6.0, 2.5.0, 2.4.3, 2.4.0, 2.15.0, 2.14.0						
	* 위 오픈소스들에 대한 최적의 조합을 정하여 빌트인 이미지형태로 탑재되어 제공 이외에도 다양한 버전별 라이브러리, 프레임워크, 개발언어 제공 (R Studio, SSH, x-window, xGBoost, PaddlePaddle,MXnet 등)							
	주요특징	※ 데이터 구축/처리, AI모델 개발, AI모델 배포/운영, HPC 인프라 관리 등 각 AI개발 단계별 최적의 환경 제공 - AI 개발 환경 : 빌트인 이미지 형태로 탑재되어 있어 선택만으로 개발환경 구성 - 데이터 셋 : 오브젝트 스토리지 기반 데이터 저장 및 관리 - 데이터 전처리 : 데이터에 대한 어노테이션 및 라벨링 제공 - 모델 서빙 : NVIDIA Triton 서버 기반 GPU 최적화된 환경에서 모델 배포 - 워크플로우 : API를 Mesh-up하고 Low-Code 기반으로 서비스 배포가 가능한 Workflow 기능 제공 - 잡스케줄링 : 모델 학습/훈련/배포 과정 효율화 및 멀티 노드 학습 지원 * 플랫폼 활용을 위한 실무교육 진행 (온라인 튜토리얼 제공, 오프라인 실무교육 진행) * 개발프레임워크, 개발언어 등 제공 버전에 따른 환경 최적화는 이용자가 직접 수행						

구분	상세	
PaaS Portal	주요특징	<ul style="list-style-type: none"> - 애플리케이션 빌드, 테스트, 배포 과정을 자동화하는 CI/CD 파이프라인 구성 - 작성한 코드 및 어플리케이션을 템플릿 기반으로 저장하고 사용하는 기능 제공. - Web IDE 기반 유연한 개발 환경 지원 - Low-Code 기반으로 애플리케이션을 구성할 수 있는 환경 제공 - 오토스케일링을 통해 자원 최적화된 환경에서 안정적인 애플리케이션 운영 - 스케줄링 기능을 통해 반복작업 최소화 환경 제공
시험형 AI 모델 (API 제공)		7종의 AI 모델 제공 <ul style="list-style-type: none"> - Gemma2 : 텍스트 임베딩 - Sentence BERT : 유사도 측정, 토픽 모델링 - Class-based Term Frequency-Inverse Document Frequency : 유사도 측정, 토픽 모델링 - Yolo V5 : 객체 탐지 - Inception V3 : 이미지 분류 - XGBoost : 데이터 분류, 주성분 분석 - LightGBM : 데이터 분류, 주성분 분석

○ 고성능 컴퓨팅 자원 지원 규모

- (지원규모) 20개 내 기업당 GPU 1식

※ 1개 사용자(기업) 당 고성능 컴퓨팅 자원 규격 (성능수치는 FP32기준의 이론적 최대 성능)

구분	GPU 종료	수량	개발 TFLOPS (FP32)	총 TFLOPS	비고
1	H100 (94GB)	3장	51 TF	153 TF	전체 사용
2	H100 (47GB, MIG)	4장	25.5 TF	102 TF	MIG (Multi-Instance GPU)
3	H100 (22GB, MIG)	3장	12.75 TF	38.3 TF	MIG (Multi-Instance GPU)
4	V100 (16GB)	12장	14 TF	168 TF	전체 사용

※ 자원 미사용 및 운영현황 조사 (인터뷰, 설문, 성과조사 참여) 에 미참여 시 사업참여 제한 및 자원이회수 될 수 있음

“ 인천스타트업파크 빅데이터·AI 플랫폼 활용지원 프로그램 ” 선정기업만 신청할 수 있음

○ 빅데이터 AI 플랫폼 활용 감사 메일

- 참여기업 : “ 모리 ” 플랫폼 활용 사례

Re: [Mori] 안녕하세요. 추가자원 대여 관련해서 문의드립니다.

4대보험...명부.pdf 10-2024...지서.pdf 스크린...21.45.png 스크린...22.28.png

안녕하십니까, Mori 공동대표 **이주호**입니다.
다음과 같은 내용 정리드립니다.

1. 압도적인 비용 절감 효과:

본 Mori는 GPU자원을 지속적으로 사용하며 연구해야하는 AI 스타트업입니다.
그러나 Runpod, AWS와 같은 클라우드 서비스의 경우 GPU자원을 시간당 단위로租用하며,
Mori가 지난 6개월간 연구를 위한 GPU 자원 대여를 그대로 진행했다고 한다면 1,400만원 이상의 지원예산이 소진되었을

특히 Vram 32GB이상의 그래픽 카드를 시중에서 구매하려면 수천만원의 자금이 필요하고,
더군다나 주문한다고 한들 수개월의 대기시간이 필요하나,
당장 연구에 필요한 자원이 가능하다는 점이 AI 스타트업에게 있어서 필수적인 요소라고 판단했습니다.

덕분에 인천 스타트업파크측의 자원 대여 덕분에 해당 예산을 아꼈고,
덕분에 웹 개발 추가 인력, AI연구 인력(2명) 고용이 가능했습니다.
해당 추가 인력 고용을 통하여 아이템 개발 속도 및 연구 속도를 끌어올릴 수 있었습니다.

실제 4대보험 가입자 명부를 참고해보시면
11월 기준으로 2명의 직원(여충환, 이다솔 직원)을 추가 고용할수 있었으며,
이는 GPU자원 대여를 통한 절감이 없었으면 불가능한 고용이었습니다.

2. 뛰어난 성능

새로 도입된 그래픽 카드(H100)는 Vram 용량이 47GB으로,
대기업에서 실제 연구 및 실험 목적으로 사용하는 장비와 동일한 사양을 지녔으며,
시드 단계의 스타트업 레벨에서는 도입하기 불가능한 가격의 장비입니다.

그러나 Mori는 POOM지원 사업을 통하여 대여가 가능하였고,
이를 통하여 일반적인 시드 레벨 스타트업에서는 전혀 불가능한 기업 수준의 연구를 설계, 진행할 수 있었습니다.

더욱이, 해당 장비가 초고속 광통신에 연결되어 있는 것과, 64코어의 인텔 제온 CPU는
실제 실험 데이터 학습을 위한 전처리 및 다운로드 과정에서 매우 쾌적한 연구가 가능했습니다.

결론적으로 POOM에서 제공해주는 장비는
단순히 GPU장비를 제공해 주었다는 것에서 그치지 않고,
'실제 실험 및 연구 목적으로 활용 가능한 장비'입니다.

이는 새로 진입하는 AI 스타트업과 기존 AI기업간의 컴퓨팅 파워 격차를 줄여주는 역할뿐만 아니라,
매우 뛰어난 성능과 그에 뒷받침 되는 광통신망에 힘 입어서 개발 생산성 및 협업 효율이 크게 개선되었습니다.

실제 특허 출원번호 통지서를 보시면,
11월에 특허 출원이 진행되었는데, 만약 POOM자원이 없었을 경우,
데스크탑 기준, 해당 실험데이터의 확보 완료 시점이 내년 2월로 밀렸습니.
POOM 자원이 없었다면 각종 정부 지원사업 과제 수행 기간내에 달성하지 못할 지점였습니다.

3. 쉬운 환경 구축과 관리

리눅스 환경에서의 GPU 활용이 매우 간단했으며, 전문 개발자가 아닌 고등학생, 대학생 수준이더라도,
초기 학습(약 20~30분) 만으로 지퍼터 노트북 기반의 작업환경에 익숙해질 수 있었습니다.

지퍼터 노트북에서 리눅스 명령어를 이용하여 기존에 데스크탑에서 진행하던 프로젝트를 github를 통해서 다운받고,
이후 리눅스 버전으로 라이브러리를 설치하면, 모든 작동 환경 구축이 끝납니다.

Mori는 최근 대두되고 있는 AI학습으로 인한 저작권 문제에 대하여
AI학습 github프로젝트를 다운받아서,
널리 사용되는 학습 방법에 대한 학습 방해 관련 실험과 검증 환경을 구축하는데 단 30분이면 충분했습니다.

더욱이, 컨테이너 별로 그래픽 카드가 분리되어 있어서
컨테이너 별로 실험 데이터를 나눠서 작업 시키는 것이 가능했습니다.

실제 첨부드린 이미지를 보시면,
하나는 노이즈를 생성하는 학습 방해를 주 용도로,
나머지 하나는 노이즈를 섞은 이미지를 통해서 학습한 뒤에 학습하는 용도로
폴 스케일의 실험을 홈페이지 하나에서 모두 관리할수 있어서 매우 사용하기 쉬웠습니다.

이렇게 말하면 너무 많이 신청해서서
제가 우수수위가 밀려서 사용하기 힘들어졌건 같아서 말하기는 조금 그러데;;;
해당 POOM 지원사업을 통해서 인천 관내의 AI 스타트업들이
한번씩은 다들 꼭 써봤으면 좋겠다는 말 남기고 싶습니다.

정말로 대단했고, POOM 지원 덕분에 이런 것들이 너무나 많습니다.
이상입니다. 감사합니다.

이주호 드림

○ 빅데이터 AI 플랫폼 활용 감사 메일 - 참여기업 : “ 애나 ” 플랫폼 활용 사례

안녕하세요 강바람 과장님 및 노형준 책임님,
주식회사 애나 이상호 입니다.

이번에 부산 에코델타시티 CCTV 군중 인식 AI 모델 학습과 관련하여,
POOM 플랫폼을 활용하여 단기간에 좋은 학습 경과를 보였습니다.

기존에 요청이 왔을 때에는 학습 자원이 부족하여 제 시간 내에 학습이 어렵다고 1차적으로 응답을 드렸었는데, 때마침 POOM 플랫폼 자원을 할당받을 수 있게 되어 제시간 내에 AI 모델 재학습 및 설치까지 완료하였습니다.

finetune before result.mp4 : 이전에 범용으로 존재하던 베이스(pre-trained) AI 모델 결과
finetune after result.mp4 : POOM 플랫폼 활용하여 AI 모델 재학습한 결과

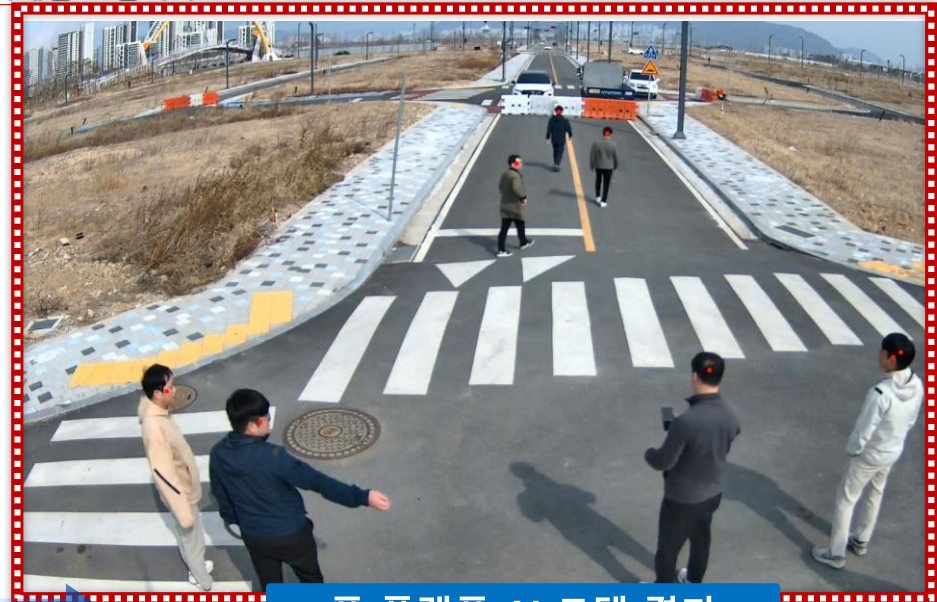
감사의 인사 및 실제 결과가 POOM 플랫폼이 있어 어떻게 가치있게 변화되었는지 공유드리고 싶어 이메일 드립니다.

참고 부탁드립니다.

감사합니다.



베이스 AI 모델 결과



폼 플랫폼 AI 모델 결과

○ 빅데이터 AI 플랫폼 및 AI 기술연계 사업 인터뷰

- 참여기업 : “ 넥스트테크 ” 의 성공 사례

Q1. 이번 인천 스타트업 AI 기술 연계 지원사업에 참여하게 된 계기는 무엇이었나요?

저희는 AI 기술이 필요했지만, 관련 인프라와 전문 인력이 부족한 상황이었습니다. 그러던 중 본 사업을 알게 되었고, AI 인프라 지원 뿐만 아니라 컨설팅까지 포함된 점에 저희에게 꼭 필요한 부분이었습니다. 특히, 아이디어를 실현 가능한 AI 모델로 구현하는 것이 가장 큰 목표였지만, AI 개발 경험이 부족하여 실제 구현 가능성, 데이터 요구사항, 모델 설계 및 개발 일정 등에 대한 명확한 가이드가 필요했습니다

Q2. ‘시각장애인 및 보행약자를 위한 횡단보도 음성 안내 시스템’을 기획한 계시는?

횡단보도에 자동 그늘막을 제작 및 설치하는 사업을 진행하면서, 시각장애인들이 횡단보도를 찾는 데 큰 어려움을 겪는 모습을 목격했습니다. 이를 계기로 시각장애인이 보다 안전하고 편리하게 길을 찾을 수 있도록 돕는 자동 안내 시스템이 필요하다고 판단했습니다. 이를 AI 기술을 활용해 시각장애인을 자동으로 인식하고, 맞춤형 음성 안내를 제공하는 시스템을 개발하면 보행 안전과 이동권 보장에 기여할 수 있을 것이라고 생각했습니다

Q3. 컨설팅의 구체적인 과정과 컨설팅을 통해 얻으신 것은 무엇인가요?

1차 컨설팅에서는 객체 탐지 기술을 활용해 시각장애인용 지팡이를 감지하고, 자동 음성 안내를 제공하는 AI 모델을 개발 목표로 설정했습니다. 이 과정에서 객체 탐지 신경망 구조, 객체 클래스 정의, 데이터 확보 및 라벨링 절차 등 핵심 기술을 학습하며, 데이터 구축 방향을 구체화할 수 있었습니다. 이후 POOM 플랫폼을 활용해 데이터 라벨링 및 모델 학습, 파인 튜닝을 진행했습니다.

2차 컨설팅에서는 모델 학습 결과를 평가하고, 감지 정확도 향상을 위한 데이터 보완 전략을 논의했습니다. 또한, 다양한 환경에서의 인식 성능을 최적화하기 위해 추가적인 학습 데이터 구성과 라벨링 기준 설정에 대한 실질적인 조언을 받을 수 있었습니다.

Q4. POOM을 활용하면서 가장 유용했던 점은 무엇인가요?

POOM 플랫폼은 고성능 GPU 자원을 제공하여 AI 개발 초기 단계에서 큰 도움이 되었습니다. 또한 API Hub를 통해 파일럿 모델을 간편하게 배포할 수 있어 실제 적용 과정이 원활했습니다. 특히, UI가 직관적으로 구성되어 있어 개발 경험이 많지 않은 팀원들도 쉽게 적용하고 데이터를 정리하고 관리하는 과정이 효율적이었으며, 라벨링을 끝낸 데이터를 손쉽게 업로드하고 학습을 진행할 수 있었습니다. 또한, 학습 도중 데이터를 수정하거나 추가하는 작업도 간편해 전체적인 작업 흐름이 원활하게 유지될 수 있었습니다.

Q5. 파일럿 모델 개발 과정과 결과는 어땠나요?

파일럿 모델 개발 목표는 시각장애인용 지팡이를 실시간으로 탐지하는 AI 모델을 구현하는 것이었습니다. 그러나 개발 과정에서 데이터 라벨링과 품질 관리가 가장 큰 과제였습니다. 초기에는 적절한 데이터를 충분히 확보하지 못했고, 경계 상자 설정 오류로 모델 학습이 원활하지 않았습니다. 이 문제를 해결하기 위해 컨설팅을 통해 CVAT 라벨링 도구 활용법과 데이터 관리 방법을 익히며 데이터 정제 과정을 개선했습니다. 특히, 경계 상자 오류에 대한 즉각적인 피드백을 받아 빠르게 수정할 수 있었던 점이 큰 도움이 되었습니다.

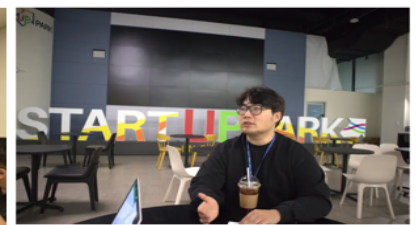
그 결과, YOLO v5 모델을 활용해 시각장애인용 지팡이를 실시간으로 탐지하는 기능을 구현했습니다. 학습된 모델은 초당 20프레임(FPS) 수준에서 탐지가 이루어졌으며, 신뢰도(Confidence) 0.85 이상을 유지했습니다. 또한, POOM 플랫폼을 활용해 학습 속도를 높이고, 컨설팅을 통해 라벨링 및 데이터 관리 방식을 익히며 모델 성능을 최적화할 수 있었습니다.

Q6. 이번에 개발된 모델의 향후 개발/활용 계획은 어떻게 되나요?

현재 개발된 파일럿 모델을 상용화하기 위한 추가 연구를 진행 중입니다. 향후에는 더 많은 데이터를 학습시켜 탐지 정확도를 높이고, 모델을 경량화하여 Jetson Nano와 같은 소형기에서도 실시간 탐지가 가능하도록 최적화할 계획입니다. 이를 통해 시각장애인의 보행 안전을 향상시키고, 스마트 교통 시스템 발전에도 기여하고자 합니다.



< AI 컨설팅 및 파일럿 개발 진행 >



< 넥스트테크 윤성호 대표 인터뷰 >