

AI-STACK

機器學習協作管理平台

加速AI導入，讓您每一分每一秒都榨滿GPU算力！

AI研發要進行機器學習/深度學習開發，需要IT人員協助建構環境、管理資源，卻不知道怎麼開始？無論是要自己買GPU，還是要上雲端，AI-STACK都能協助你規劃需求，輕鬆管理IT與使用者流程，馬上行動與我們聯繫吧！

1

叢集運算

- 使用Kubernetes與Docker容器化技術管理
- 可整合公有雲快速擴充GPU資源
- 可同時管理多台GPU伺服器、混GPU卡環境
- 支援多節點運算(政府版授權提供)

2

快速建立研發環境

- 簡易步驟快速建立(個人/團隊)ML環境
- 提供Jupyter Notebook、JupyterLab開發工具
- 支援NGC AI深度學習框架,支援自製鏡像

3

帳號控管、群組資源配額與權限管理

- 支援Single Sign-On(SSO),可介接LDAP/AD/OpenID/Oauth2
- 運算資源隔離,SSH安全殼層防護,可設定IP白名單
- 簽核系統可依部門層級或專案需求建群組、權限、額度

4

批次建立、資源共享、儲存服務

- 可批次建立容器,安排預約申請作業,批次任務定期排程
- 提供多人多容器建立在單一GPU卡上運行
- 支援儲存服務(eg.NFS、S3)

功能比較表

模組	功能	細項設定	Enterprise	Express
Web 介面	User Portal		●	●
機器學習服務 MLS	容器服務列表	容器資訊	●	●
		容器監控	●	●
		刪除容器	●	●
機器學習服務 MLS 快速容器服務 RCS	新增容器服務	新增容器	●	●
		自選 AI framework	●	●
		批次建立容器	●	●
	AI 訓練任務列表 自定義鏡像	AI 訓練任務資訊	●	●
		刪除 AI 訓練	●	●
		新增自訂義鏡像	●	●
	AI 訓練任務列表 共享 GPU	自訂義鏡像列表	●	●
		提供多人多容器建立在單一 GPU 卡上運行	●	●
	GPU 使用時數 儲存管理	統計容器內真實 GPU 使用時數	●	●
		建立與管理容器所使用的資料儲存空間	●	●
部署管理		●		
ConfigMap 管理		●		
網路政策管理		●		
快速容器服務 RCS 模型推論部署服務	資源監控		●	
	事件檢視		●	
	鏡像列表		●	
	模型列表		●	
	推論服務		●	

* AI-Stack Express 基本套件包含身份認證套件 (連接 LDAP)

* AI-Stack Enterprise 基本套件除了身份認證套件 (連接 LDAP), 還包含 NAS 整合連接