AI/ML Platform

Professor：Jenq-Shiou Leu
Student：Antony Wang
Date：2023/08/08
National Taiwan University of Science and Technology

MITLab©
Mobilizing Information Technology Laboratory

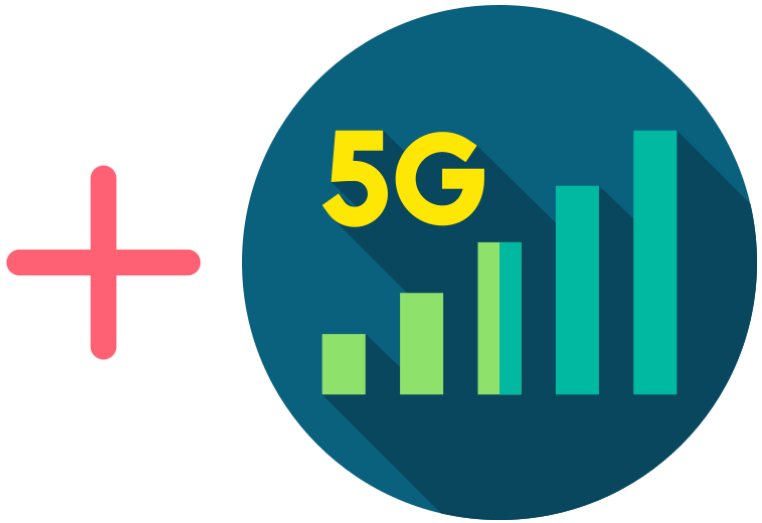
Outline

- Introduction
- Model Training Platform
- Model Training Process
- OSC release-I
- AI/ML Workflow
- Discussion

Introduction

An AI/ML platform combined with 5G O-RAN provides regular model training and evaluation.

AI/ML



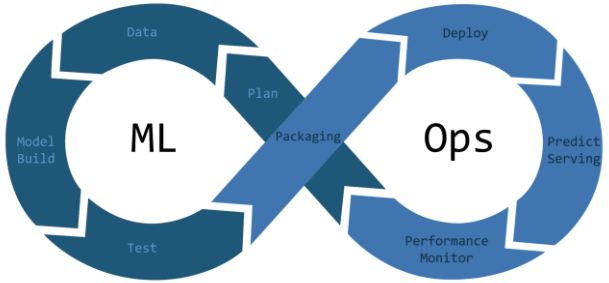
5G O-RAN

Open Source

Model Training Platform

MLOps: Machine Learning Operations

- **Plan → Data → Model Build → Test → Packaging → Deploy → Predict Serving → Performance Monitor**



https://coffeewhale.com/what-is-mlops

Tools

- **Kubernetes**
- **Kubeflow**
- **KServe**

kubernetes

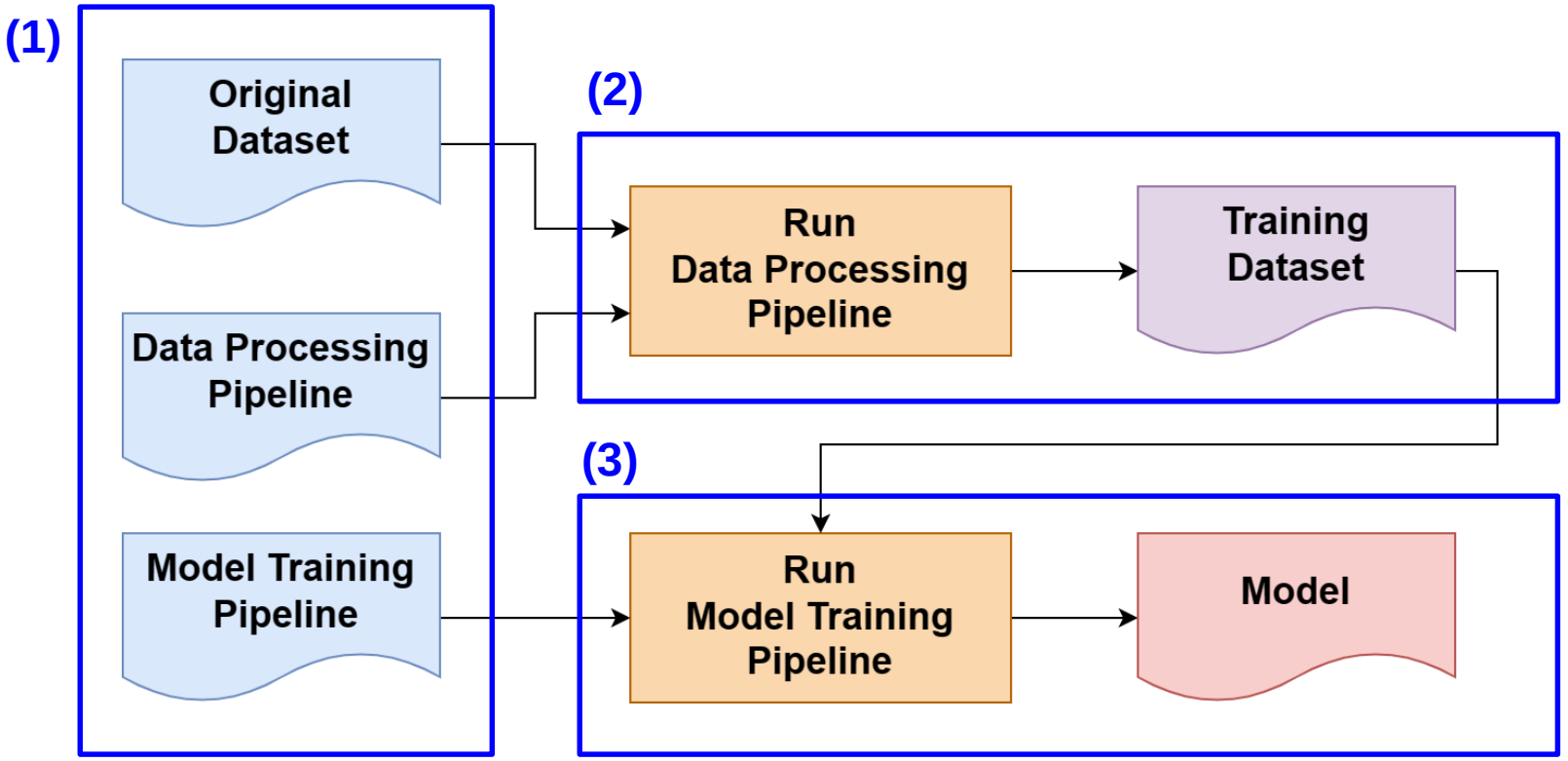
Kubeflow

KServe

Model Training Platform

- **Model Pipeline Training & Experimentation**
- **Distributed Model Training Resource Management**
- **Model Training Management & Monitoring**

Model Training Process with Pipeline



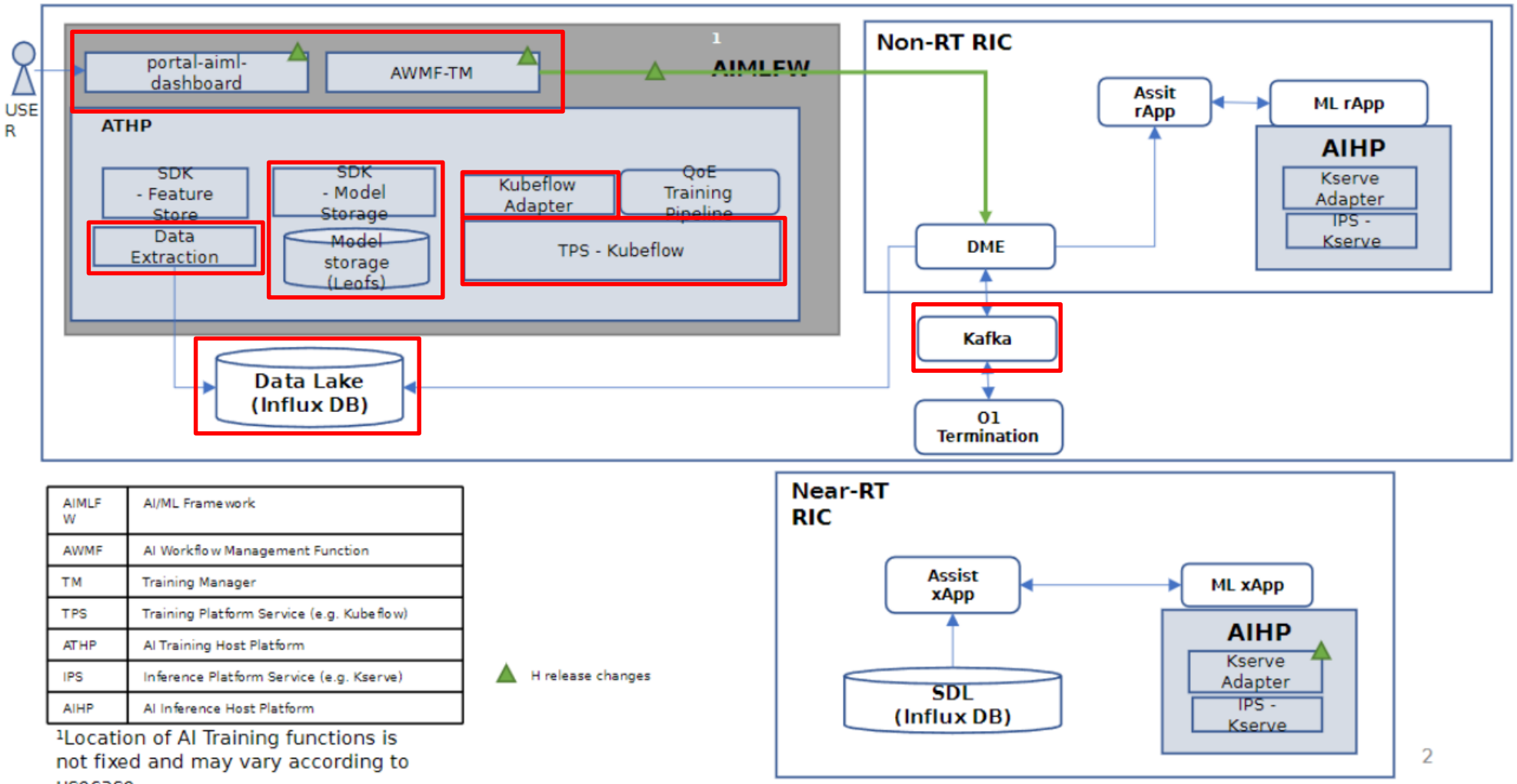
(1) Prepare **Original Dataset** & **Pipeline**

(2) Run Data Processing Pipeline → **Training Dataset**

(3) Run Model Training Pipeline → **Model**

OSC Release-I

: We have impelement

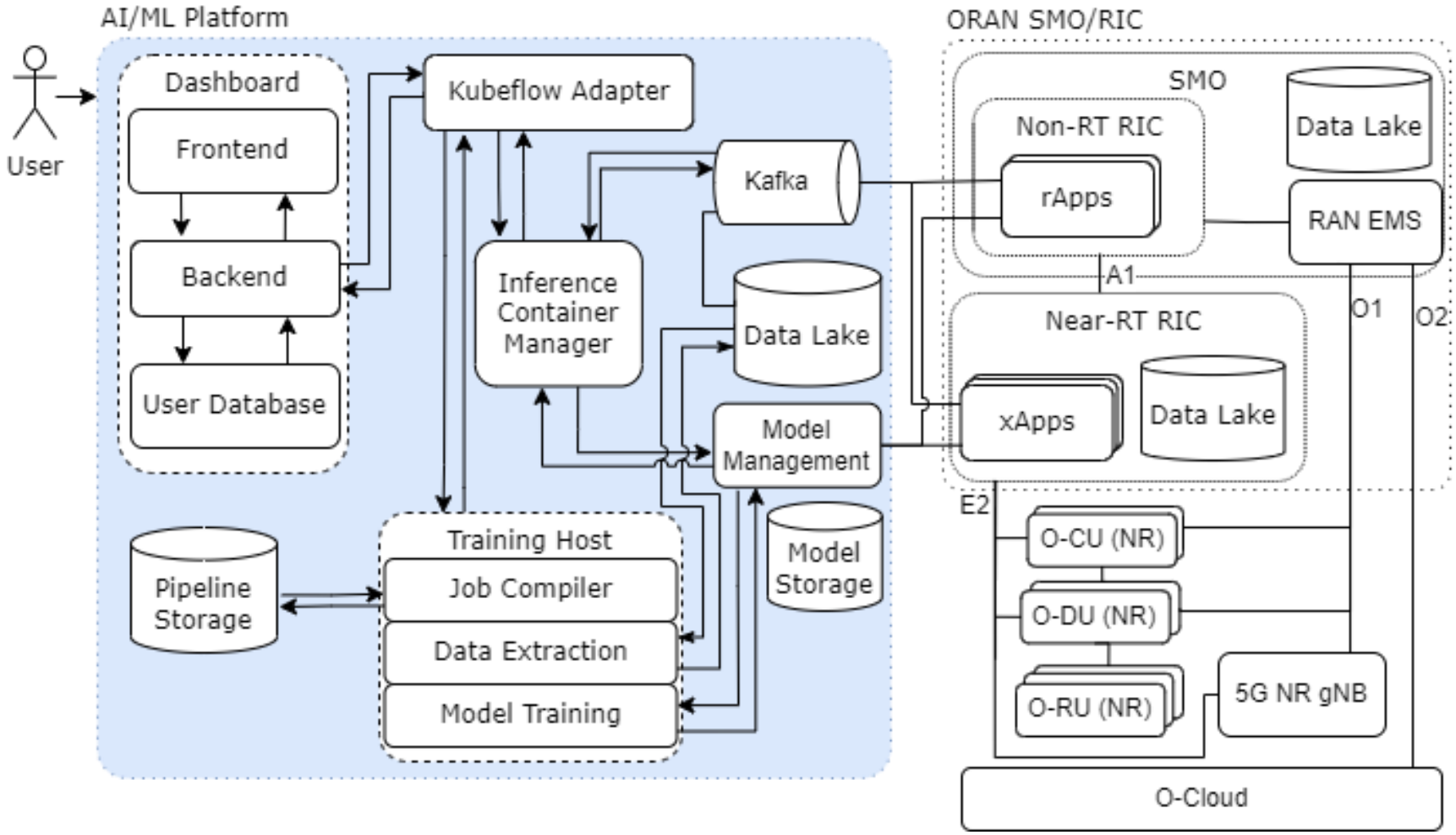


AIMLFW	AI/ML Framework
AWMF	AI Workflow Management Function
TM	Training Manager
TPS	Training Platform Service (e.g. Kubeflow)
ATHP	AI Training Host Platform
IPS	Inference Platform Service (e.g. Kserve)
AIHP	AI Inference Host Platform

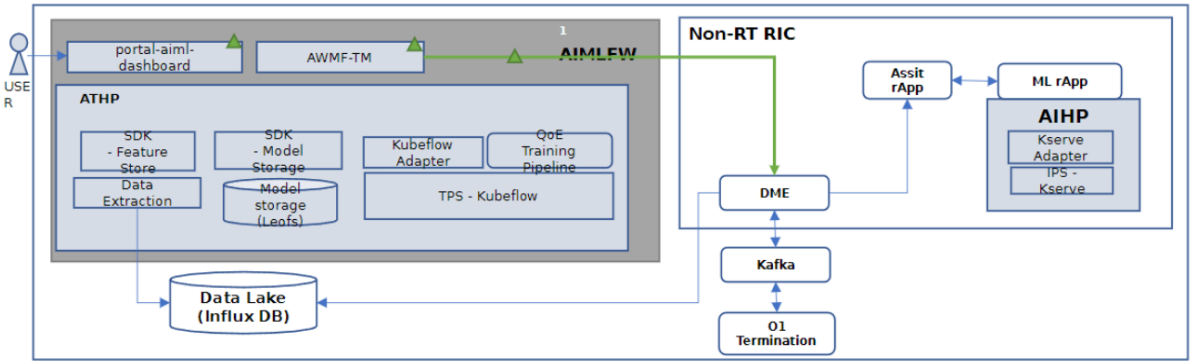
H release changes

[1]Location of AI Training functions is not fixed and may vary according to usecase.

AI/ML Workflow



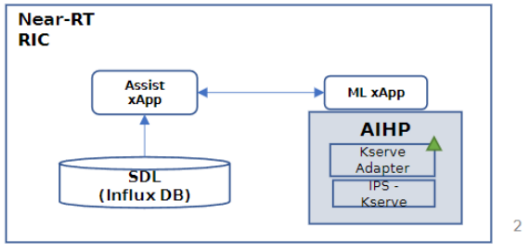
Discussion



AIMLFW	AI/ML Framework
AWMF	AI Workflow Management Function
TM	Training Manager
TPS	Training Platform Service (e.g. Kubeflow)
ATHP	AI Training Host Platform
IPS	Inference Platform Service (e.g. Kserve)
AIHP	AI Inference Host Platform

▲ H release changes

[1]Location of AI Training functions is not fixed and may vary according to usecase.



Generic Training Pipeline

- Generic Training Pipeline not specific to a usecase
  - Provide features, hyperparameters required for training as input
- Need metadata stored regarding the pipeline
  - Pipeline hyperparameters
  - Type of algorithm
  - Libraries used
  - Model format
- Need to check if Kubeflow feature can be used for storing and retrieving metadata. Another option is to store in training manager.
- The pipeline metadata will help to select right pipeline in training when required information is provided. Currently we manually select from GUI.
- Option to provide the pipeline during installation and also dynamically after installation

- Services and APIs to be planned based on Non-RT RIC specifications

Service	Procedure/API	Details
Model management and exposure (MME) service	• Register AI/ML model • Deregister AI/ML model • AI/ML Model Update • Store AI/ML model • Delete AI/ML model • Discover AI/ML model • Subscribe AI/ML models changes • Unsubscribe AI/ML models changes • Notify AI/ML models changes • Register AI/ML model training capability • Deregister AI/ML model training capability	
Training Service	• Request AI/ML training • Query AI/ML training job status • Cancel AI/ML training • Notify AI/ML training job status change	

Q1. What is the feature store? What is inside it? Why called SDK?

Q2. How does the API update the job status and notify if the job status is changed, and what protocol is adopted?

Q3. Can the general training pipeline do the training if the data source or characteristic is different?

Q4. What is the purpose to have an automated testing for AI/ML Framework?