



iImagine Competence Center: Facilitate collaboration between Scientists

Fahimeh Alibabaei, Elnaz Azmi, Valentin Kozlov

but also UPV, IFCA, EGI, ...

Dissemination level: Public

02–06 June 2025

EGI2025, Santander, Spain

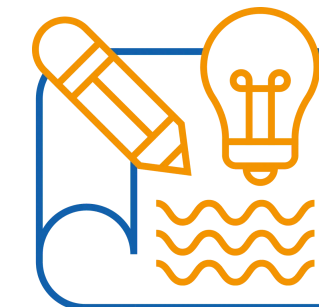
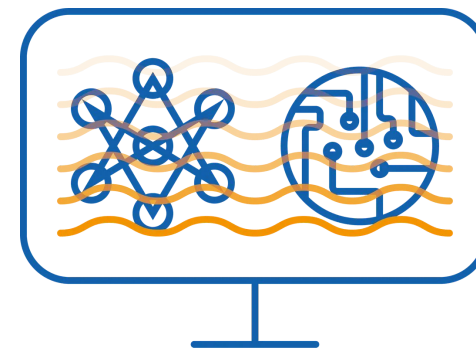


Funded by
the European Union

iImagine receives funding from the European Union's Horizon Europe
research and innovation program under grant agreement No. 101058625

Outline

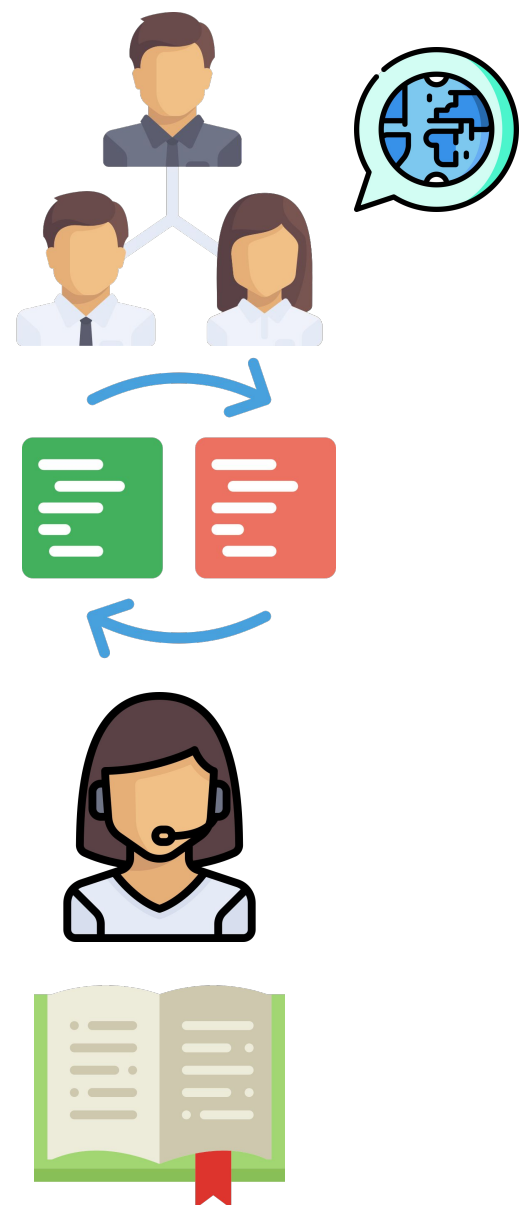
- iImagine Competence Center
- Use Cases
- Best Practices
- Summary





Mission: Facilitate the path from image **data** to imaging **applications** and **services** for scientists in aquatic sciences

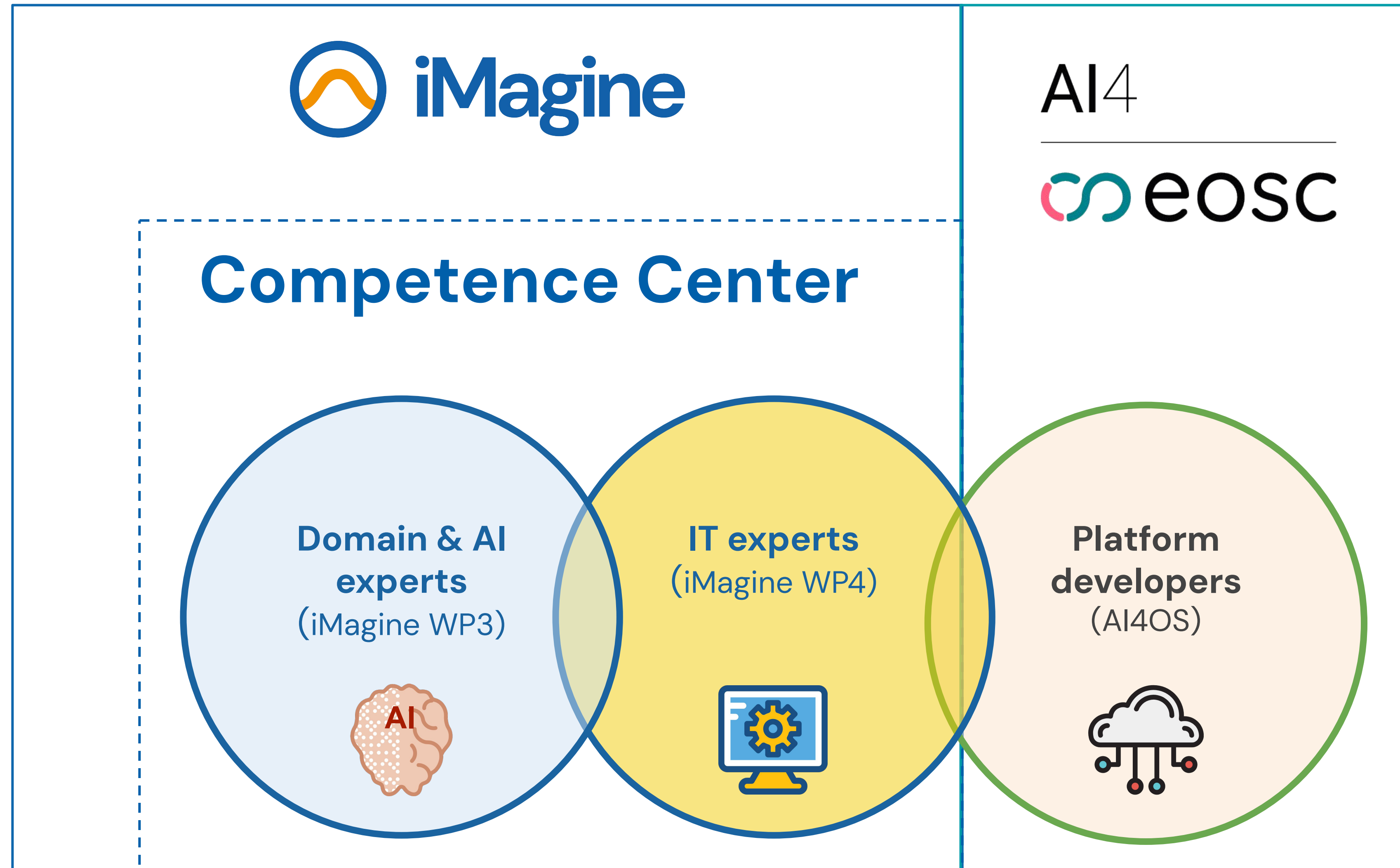
Methodology:



- Connect **experts** from different domains (science, AI, IT, ...)
- Stimulate **knowledge exchange** between the parties
- Speak **common language**
- Provide **training and support**
(data, AI, FAIR, platform, service management)
- Create **best practices**

Competence Center:

Synergetic approach



Competence Center:

Support

Knowledge exchange:

- Regular meetings with iImagine and Open Call use cases
- We try to understand each other & speak **same** language

Training and Support:

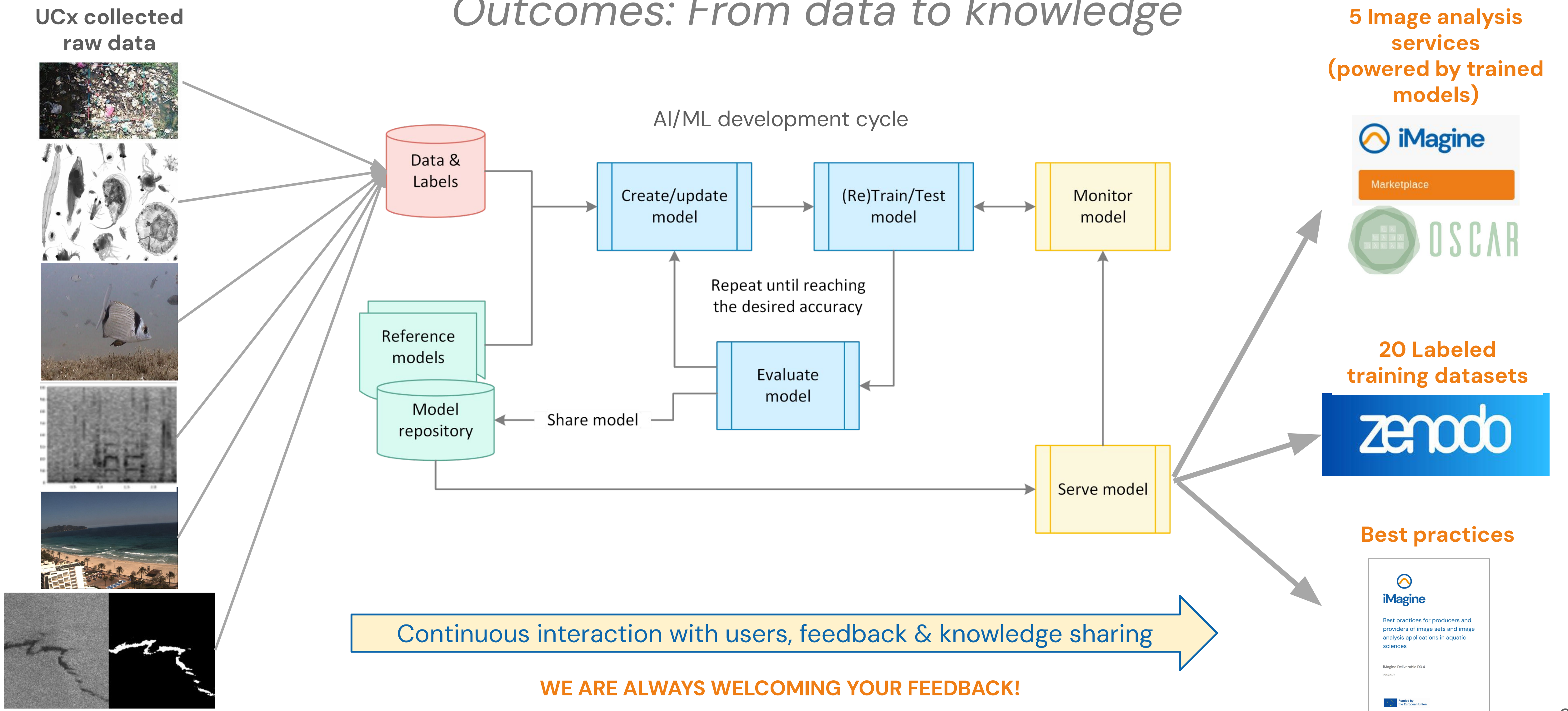
- Competence center workshops ([2023](#), [2024](#), [2025](#))
- Dedicated [events](#) and [webinars](#)
- AI4EOSC platform [user workshop](#) (2023) and [webinars](#)
- [FitSM](#) course for imagine service providers (EGI)
- [Support email list](#), individual meetings and consultancy, e.g.
 - ▷ Marketplace publishing
 - ▷ Infrastructure (e.g. OSCAR, thanks to UPV!)


Best practices: [documentation](#), [deliverables](#), articles, [datasets](#)



Competence Center:

Outcomes: From data to knowledge





Dashboard

Statistics

Catalog

Modules

Tools

Tasks

Try me

Deployments

Inference



Platform Services

Storage

Experiment tracking

AI Inference workflows

The iImagine dashboard is a service provided by CSIC, co-funded by [iImagine](#)



Powered by [AI4OS](#) [AI4](#) [eosc](#)

Libraries

Tasks

Platform Categories

Data Types

Tags

Add filter

Cold coral segmentation

Cold corals segmentation model

AUC1

Conus species classifier

Classify conus images among 70 species.

Deep Species Detection

Deep Sea Species Detection

UC3a

Diatom detection

Diatoms detection with oriented bounding boxes.

Litter assesment

AI-powered plastic litter counter, from drone footage to cleaner waters.

UC1

Object Detection and Classification with Pytorch

A trained Region Convolutional Neural Network (Faster RCNN) for object detection and classification.

iImagine

Object detection with FasterRCNN

Object detection using FasterRCNN model(s)

iImagine

OBSEA Fish Detection

AI-based fish detection and classification algorithm based on YOLOv8.

UC3o

Phytoplankton species classifier (VLIZ)

Identify the species level of Plankton for 95 classes.

UC5

smartbay-marine-types

AI-based Marine animal types detection and classification algorithm based on YOLOv8. The model has been tuned to detect and classify marine animal types at the Smartbay underwater observatory

UC3s

smartbay-prawn-burrow-detection

AI-based Prawn Burrow detection and classification algorithm based on YOLOv8. The model has been tuned to detect and classify Prawn Burrows in Irish Prawn (Nephrops norvegicus) Fishery Areas.

UC3s

smartbay-species-detection

AI-based Marine Species detection and classification algorithm based on YOLOv8. The model has been tuned to detect and classify marine species at the Smartbay underwater observatory.

Train an audio classifier

Train your own audio classifier with your custom dataset. It comes also pretrained on the 527 AudioSet classes.

iImagine

Train an image classifier

Train your own image classifier with your custom dataset. It comes also pretrained on the 1K ImageNet classes.

iImagine

Train an underwater-noise-classification (VLIZ)

Train your own underwater-noise-classifier with your custom dataset to identify ships

UC6

WITOIL-for-iImagine

Medsluk-II Oil fate lagrangian model

UC4

YoloV8 model

Object detection using YoloV8 model

iImagine

ZooProcess Multiple Classifier

A module to differentiate images containing multiple zooplankton objects from those containing only one object

UC2

ZooProcess Multiple Separator

A module to draw lines between objects to be separated from one another in ZooScan images

UC2

iImagine Services Access Request

[iImagine](#) has released its imaging analysis services for aquatic sciences. Be the first to request access and start using them.

[In Google anmelden](#), um den Fortschritt zu speichern. [Weitere Informationen](#)

** Gibt eine erforderliche Frage an*

E-Mail-Adresse *

Ihre E-Mail-Adresse

Your Name *

Meine Antwort

Your Affiliation *

Meine Antwort

I work in a project and want to get access to one or more services (use the text box below to give us more details!)

Meine Antwort

Select the Service You Want to Have Access *

☐ Litter Assessment: Identify Floating Plastic for a Cleaner Future

☐ ZooProcess: Unveiling the Secrets of Plankton

☐ AI-Powered Fish Monitoring for Healthy Oceans: EMSO-OBSEA

☐ AI-Powered Fish Monitoring for Healthy Oceans: EMSO-Azores

☐ AI-Powered Fish Monitoring for Healthy Oceans: EMSO-SmartBay

☐ WITOil: Predicting Oil Spills for Cleaner Seas

☐ Uncovering Phytoplankton with AI

Select the Type of Access You're Interested In *

☐ Analyse: Analyse Your Whole Image Set With the Analysis Service

☐ Re-training - Re-train the image analysis model with your training images to reach more precise predictions for your use case.

☐ Deploy - Download the trained image analysis model and use it on local computing resources - either for image analysis or for re-training.

How to Access

Try out the AI modules on the iImagine AI Platform for 10 minutes

- Check this [tutorial](#) first
- Log in on the [iImagine AI dashboard](#) with any account
- Select the module of interest
- Click on the TRY button on the right. A Gradio interface will open.

Request Access

Fill in the form [here](#).

Thematic Services for Image Analysis

<https://www.imagine-ai.eu/services/image-analysis-services-for-aquatic-sciences>



WE ARE OPEN FOR
BUSINESS!

Best Practices:

Publication

- Lessons learned during the whole AI development cycle
- Experiences from 8 aquatic image analysis use cases
- Available since October 2024 at Zenodo:
<https://doi.org/10.5281/zenodo.13864196>
- Accepted for publication at the “Ecological Informatics” peer-reviewed journal




iImagine


Best practices for producers and providers of image sets and image analysis applications in aquatic sciences

iImagine Deliverable D3.4

01/10/2024

**Funded by the European Union**

iImagine receives funding from the European Union's Horizon Europe research and innovation program under grant agreement No. 101058625. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union, which cannot be held responsible for them.

**iImagine project**
Part of EU Open Research Repository


Published October 1, 2024 | Version 1.1 Final

Project deliverableOpen

iImagine D3.4 Best practices for producers and providers of image sets and image analysis applications in aquatic sciences

Alibabaei, Khadijeh¹ ; Azmi, Elnaz¹ ; Kozlov, Valentin¹

Hide affiliations

1.  Karlsruhe Institute of Technology

iImagine is a European project to serve aquatic researchers with a suite of high-performance image analysis tools equipped with artificial intelligence. To effectively achieve the objectives of the project, eight use cases in different areas of aquatic science are collaborating with the providers of the iImagine AI platform. This collaboration has yielded valuable insights and practical knowledge.

In this deliverable, we delve into the details of the best AI-based solutions for image processing in aquatic science, drawing on the extensive experience and knowledge we have gained over the course of the iImagine project. We thoroughly review the methods and tools used in the initial phase of data labelling, in the subsequent phases of model training and in the final deployment of the model as a service.

Files

iImagine D3.4 _V1.1_Final.pdf

241
VIEWS

209
DOWNLOADS

Show more details

Versions

Version 1.1 Final	Oct 1, 2024
10.5281/zenodo.14893997	
Version 1.1 Under EC Review	Oct 1, 2024
10.5281/zenodo.13866394	
Version Under EC Review	Oct 1, 2024
10.5281/zenodo.13864197	

View all 3 versions

Cite all versions?

You can cite all versions by using the DOI [10.5281/zenodo.13864196](https://doi.org/10.5281/zenodo.13864196). This DOI represents all versions, and will always resolve to the latest one. [Read more.](#)

Best Practices:

Covered topics



Deep Learning

Annotation Tools

Data Repositories and Open-source Dataset for Marine Applications

Preprocessing Techniques

Performance Metrics and Evaluation Methods

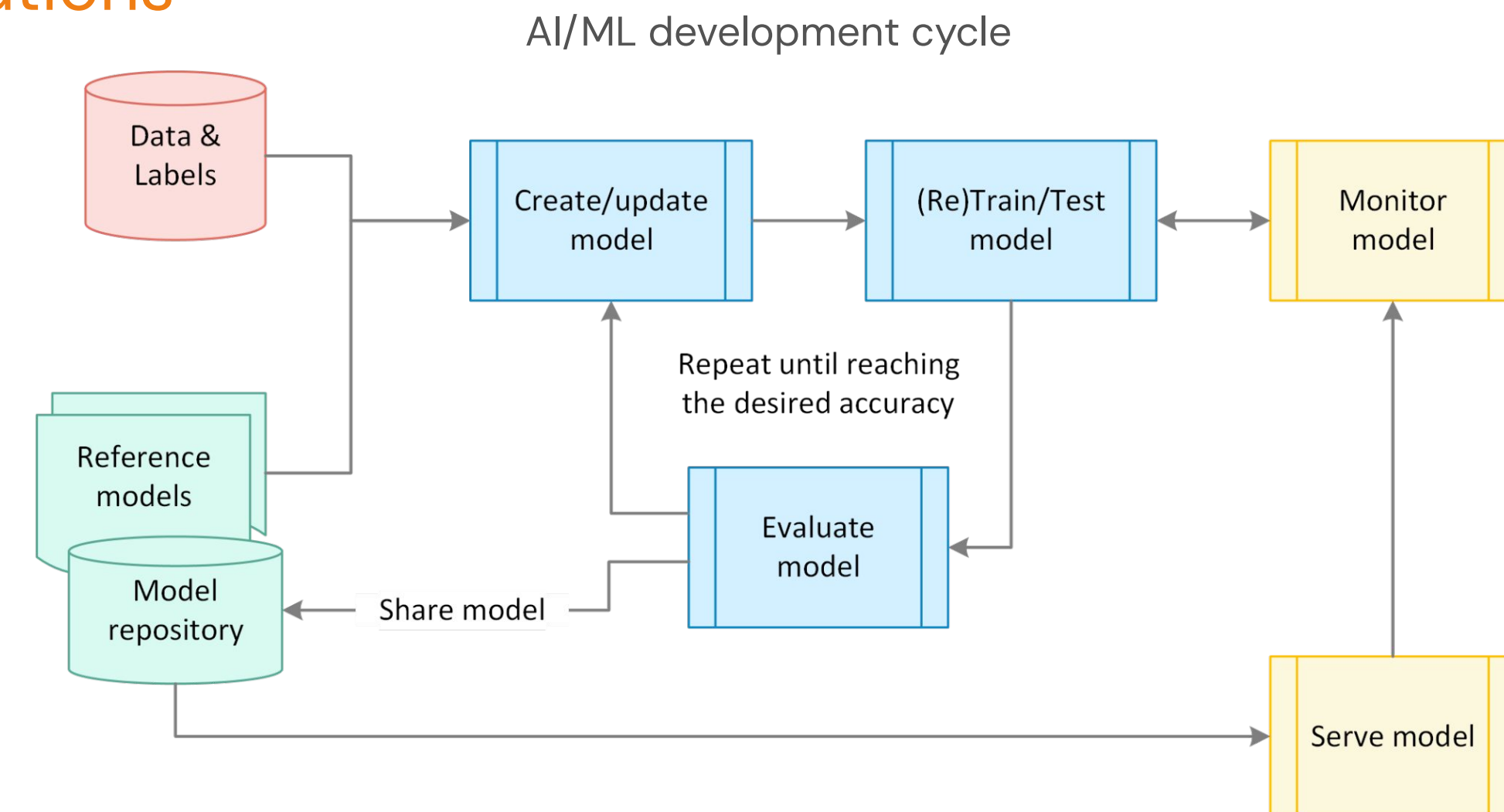
Tools for Monitoring Model Performance

Data Biases and Fairness in Aquatic Science Models and Data

Model Delivery

AI Model Drift Tools

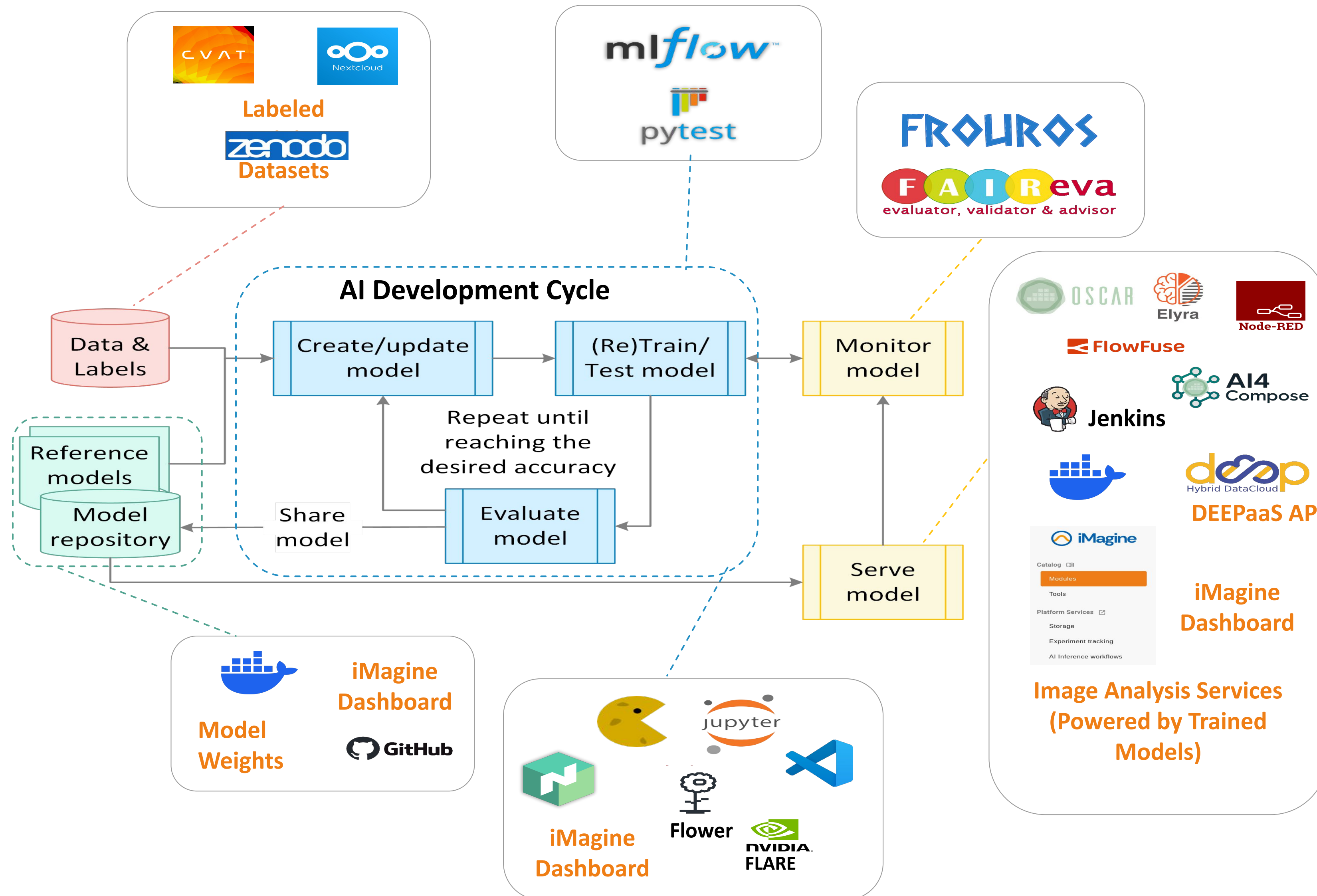
Use Case Experiences



Diversity in AI development

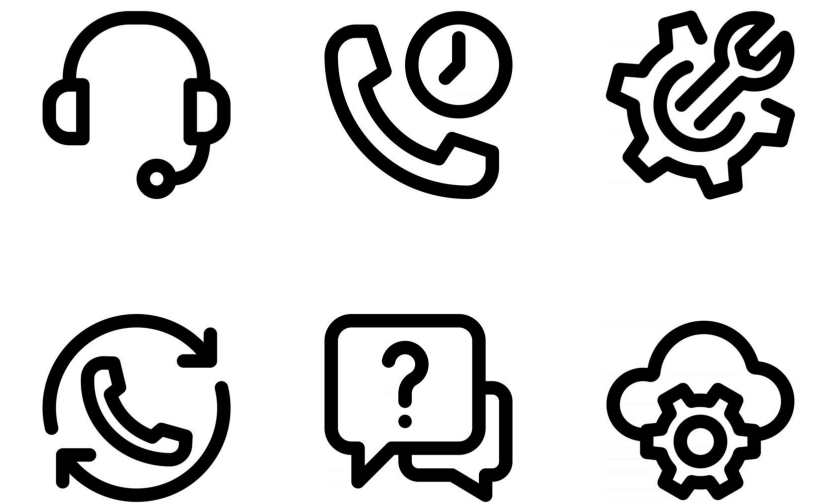
Use cases	Labeling	Preprocessing	Data biases	Models	Performance metrics
UC1 – Marine litter assessment	prelabeled with in-house litter categories(JLIST)	resizing, augmentation , (horizontal flips, vertical flips, and random rotations)	unbalanced dataset, AI model checked, no need to address	MobileNetV2 , SqueezeNet1.1., DenseNet121 , ResNet50	accuracy, precision, recall, F1-Score
UC2 – ZooScanNet	labeling with EcoTaxa, segmentation with current version of ZooProcess (both in-house tools)	Before AI: background subtraction on the original image (with a blank scan, by ZooProcess) Classif: data augmentation for (rotation, resizing, cropping, flipping) done just-in-time during training and inference Segment: no augmentation	unbalanced dataset, combination of resampling (for classif and segment) and weights (for classif)	Classif: MobileNet Segment: MaskRCNN, Mask2Former	Classif: binary cross-entropy Segment: accuracy, recall, precision, IoU
UC3o – EMSO OBSEA	labelling , roboflow, label studio, BIIGLE	augmentation	cover the seasonality , one-year dataset, unbalanced dataset, adding dataset of other species	Faster-RCNN, YOLOv8	mAP@50, precision, recall
UC3a – EMSO Azores	annotated by citizens	padding, bounding box corrections, addressing redundancy	unbalanced dataset, focused on two specific species	YOLOv8	precision, recall , mAP50–95
UC3s –Smartbay	annotated by CVAT	augmentation (roboflow noise and contrast)	different sediment types, larger training datasets and more classes	YOLOv8 , DOVER VQA	mAP50–95
UC4 – Oil spill detection	proprietary method developed by Orbital EOS	weight-loss	false positives, almost 98% of pixels are not real oil spills	Bayesian optimization	Fractions Skills Score
UC5 – Flowcam plankton identification	in-house developed labeling tool	augmentation	unbalanced dataset, resampling	CNN based on an Xception architecture	precision, recall, F1 score
UC6 – Underwater Noise Identification		window size, features extraction ,	data from different locations and dates during one year	Contrastive Language-Audio model based on the CLAP model	RMSE
UC7 – Shoreline extraction, Beach monitoring	CVAT , in-house tool written in R	label format adaptation, colour space transformation, data scaling , split rows, patchify, augmentation	unbalanced dataset, (geographic locations and/or camera FOVs)	U-Net, YOLOv8 ,9,11-seg, Bi-LSTM	accuracy , loss, precision, recall, F1 score, IoU
UC8 – Freshwater diatoms identification	BIIGLE , LabelBox	pre-training on synthetic datasets, padding, bounding box corrections (resizing, rotation), features extraction	unbalanced dataset	YOLOv5 , EfficientNet, MAPLE, SAM2	accuracy , Expected Calibration Error, Negative Log-Likelihood, Precision-Recall curve

Tools and Infrastructures



Summary

- The Competence Center of iImagine facilitates the transition from image data to imaging services for scientists in aquatic sciences through various means.



- **Key Outcomes of Collaboration within the Competence Center:**

- Collaboration between researchers, developers, and infrastructure providers
- AI-powered image analysis services
- Support and training for scientists
- Best practices for the aquatic science community



[Marketplace](#)



[Best Practices](#)



[imagine-ai.eu](#)



[twitter.com](#)



[zenodo.org](#)



Thank you for your attention

imagine-ai-platform-support@mailman.egi.eu

imagine-ai.eu



Funded by
the European Union

iMagine receives funding from the European Union's Horizon Europe
research and innovation program under grant agreement No. 101058625

- CVAT: user-friendly interface, robust annotation features

Use Cases Overview

5 mature + 3 prototypes

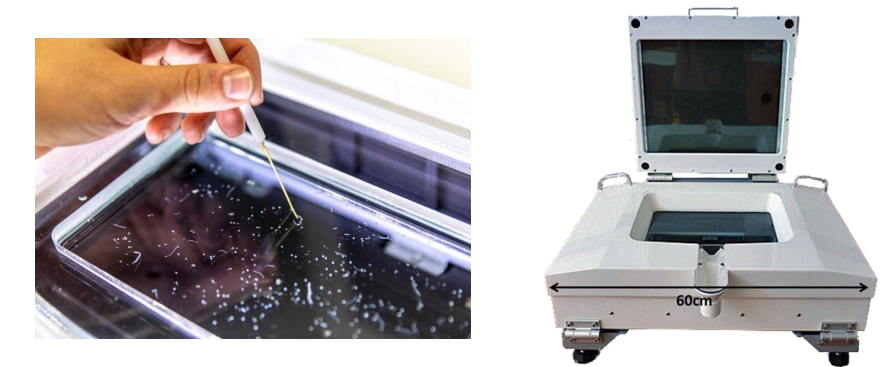
Aquatic Litter Drones

(DFKI, MARIS, OGS)



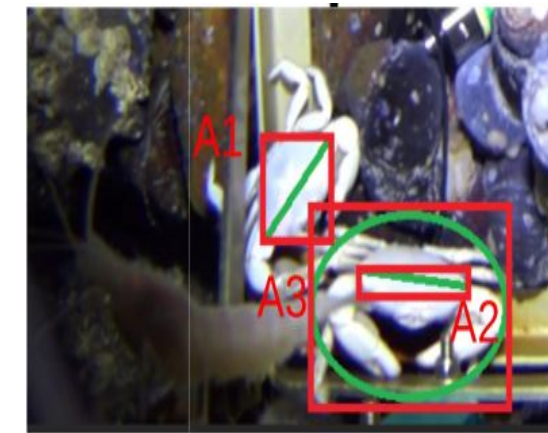
Zooscan – EcoTaxa pipeline

(Sorbonne Uni.)



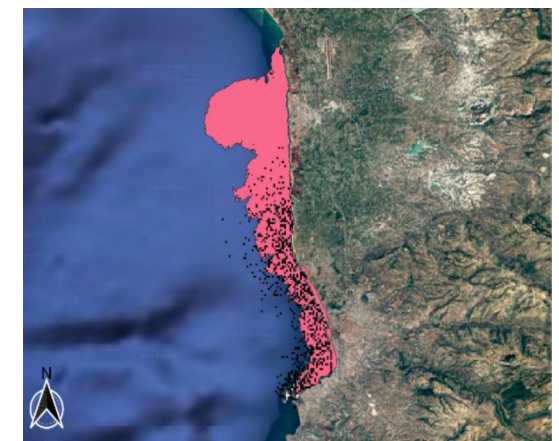
Marine Ecosystem Monitoring

(EMSO ERIC, UPC, IFREMER, MI)



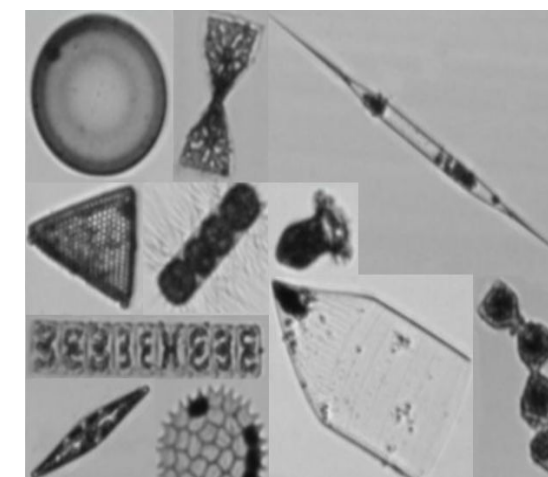
Oil Spill Detection

(CMCC, OrbitalEOS, UNITN)



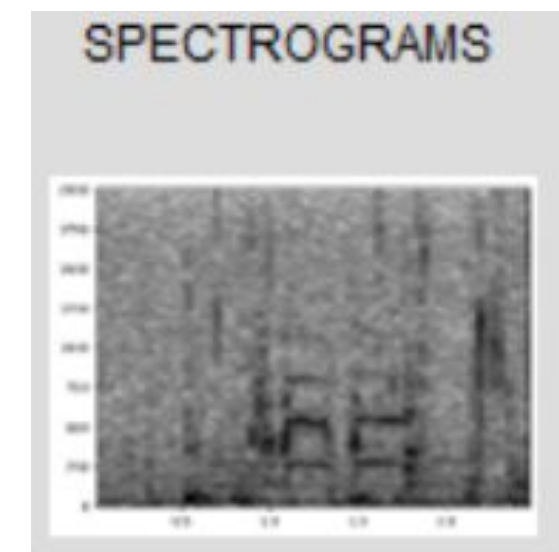
Flowcam phytoplankton

identification (VLIZ)



Underwater noise identification

(VLIZ)

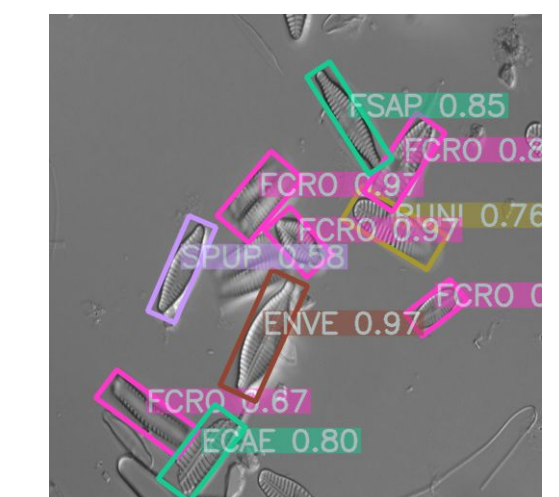


Beach monitoring (SOCIB)



Freshwater diatoms

identification (UL-LIEC)

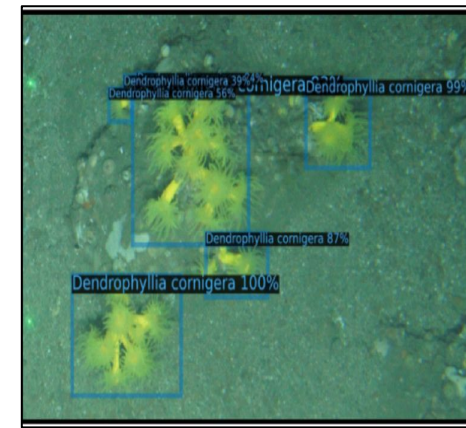


Open Call Use Cases

Onboarded via Open Calls

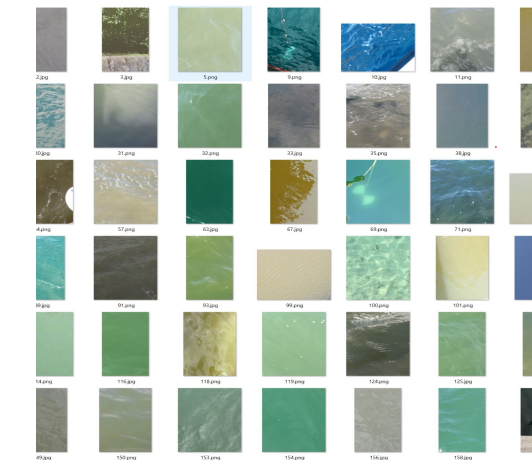
Improving knowledge about Cold Water Coral Reef

Spanish Oceanographic Institute (ES)



Validating images collected with the EyeonWater app

MARIS (NL)



Satellite-Derived Bathymetry

Institute of Marine Sciences of Andalusia (ES)



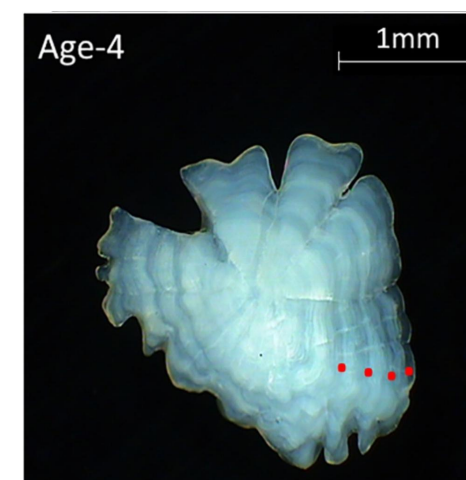
DEcentrAlised Learning (DEAL)

Plymouth Marine Laboratory, UK



AI for image based age reading from fish otoliths

Danish National Institute of Aquatic Resources (DK)



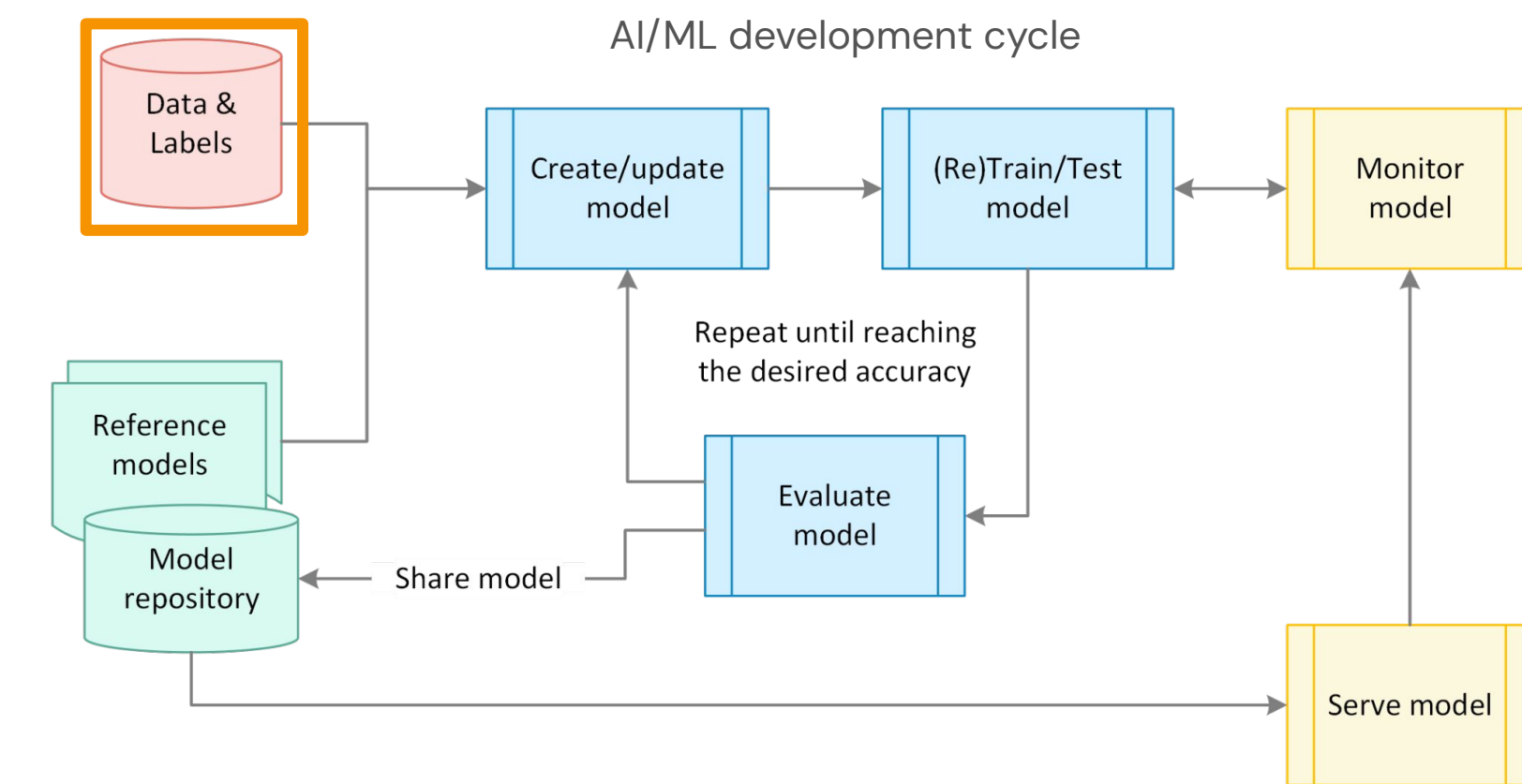
Sea Wave and coastal inundation detection Methodology (SWiM)

LNEC PT



Best Practices:

Data Repository



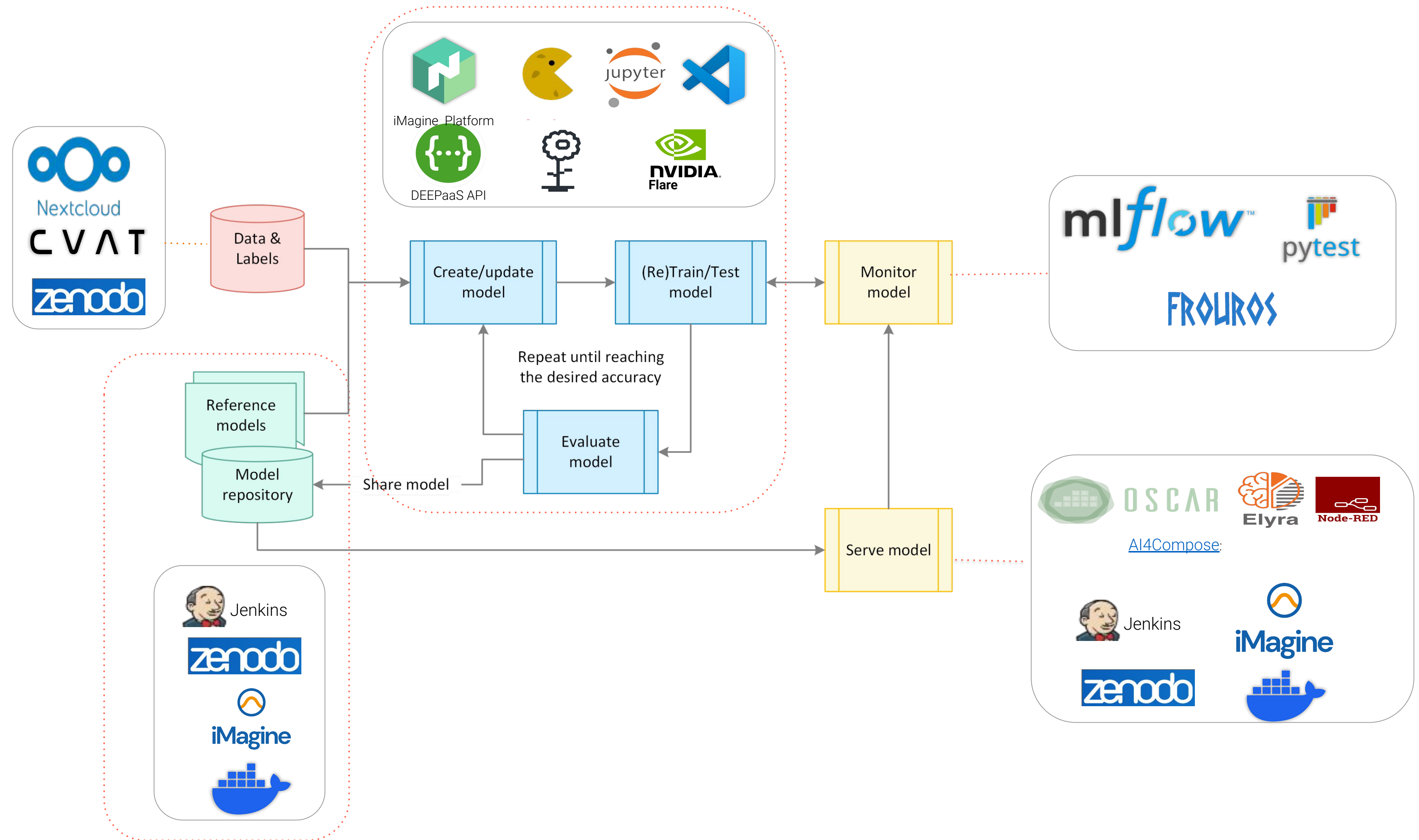
Data Repositories and Open-source Dataset for Marine Applications

- **Zenodo**: robust solution, open-access platform to store and share datasets
- iImagine training datasets are in the **iImagine community at Zenodo**



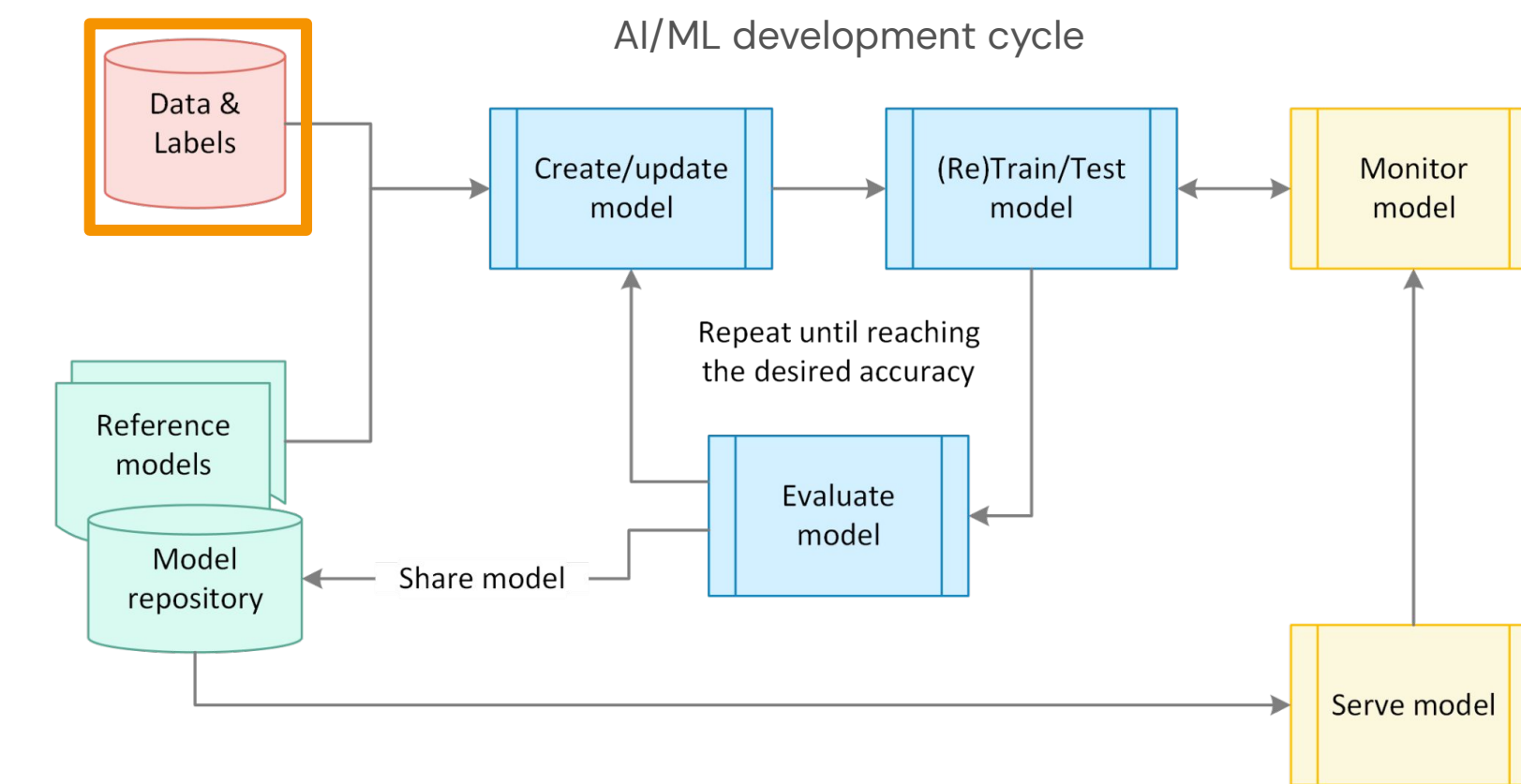
iImagine and Zenodo **extend metadata** format for aquatic science

Tools and Infrastructures



Best Practices:

Data Bias



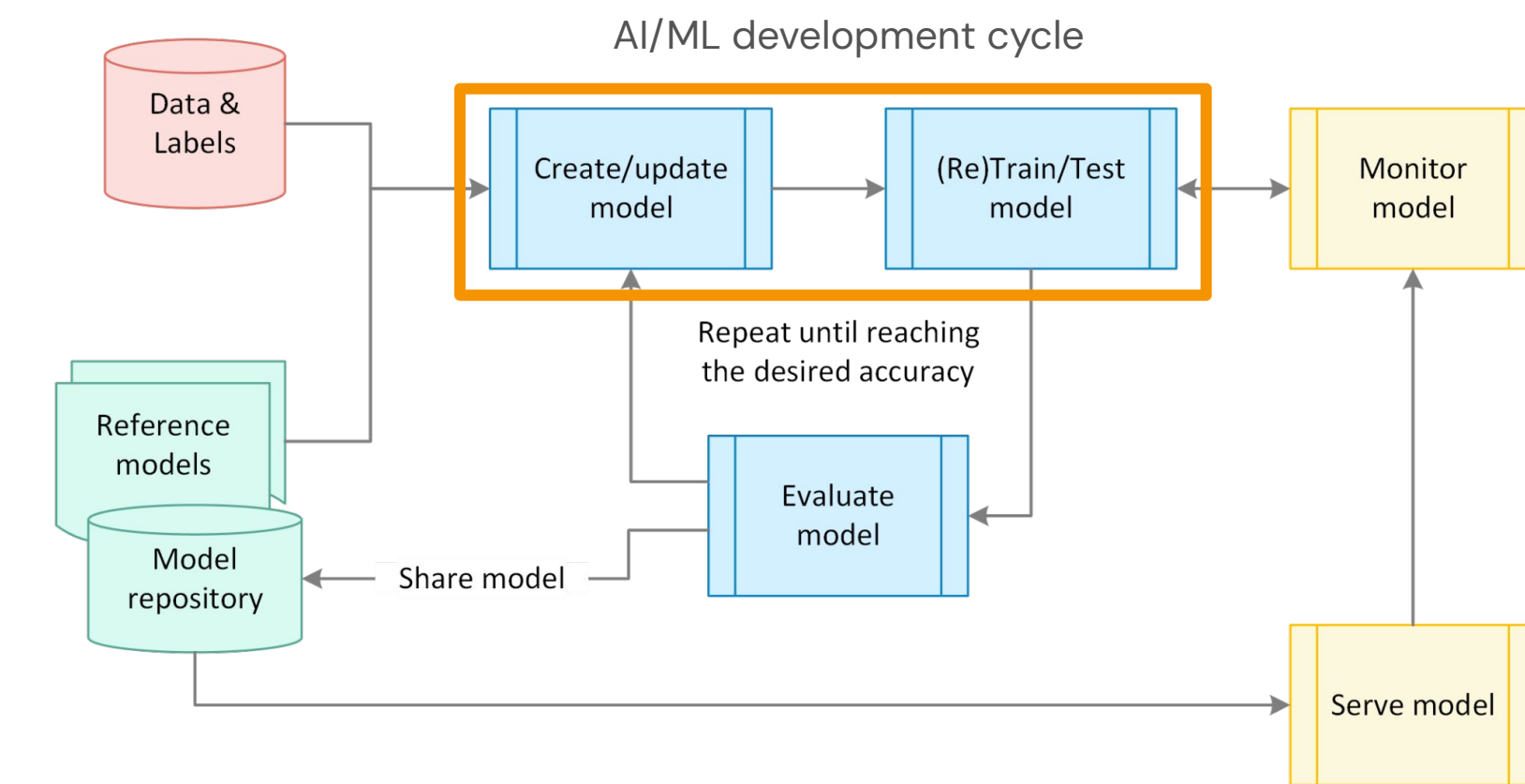
Data Bias and Fairness in Aquatic Science Data

- Recommendations about handling **imbalanced datasets**
- **FAIR EVA** tool to improve the FAIRness of published training datasets concerning the metadata used in Zenodo



Best Practices:

Create/Update/Train/Test Model



● Deep Learning Models

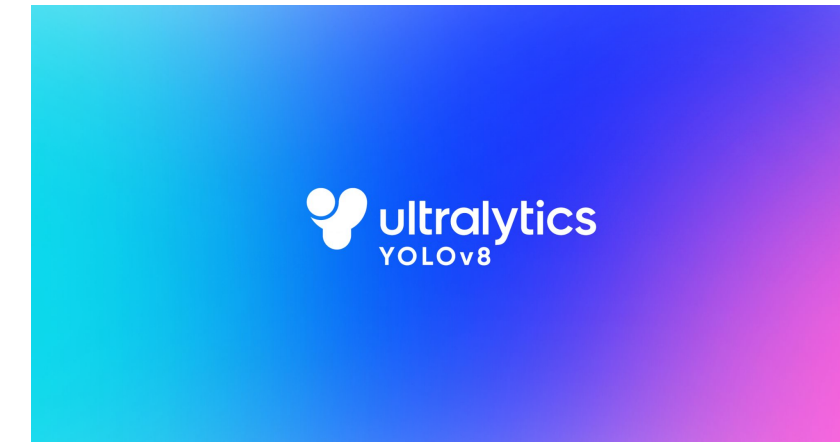
- Classification
 - Mobilenet
- Object Detection
 - YOLO: most commonly used, showed strong performance
- Segmentation
 - YOLO
 - Mask2Former: achieved better performance in panoptic segmentation than instance segmentation with Mask R-CNN

● Training and Testing on iImagine Marketplace

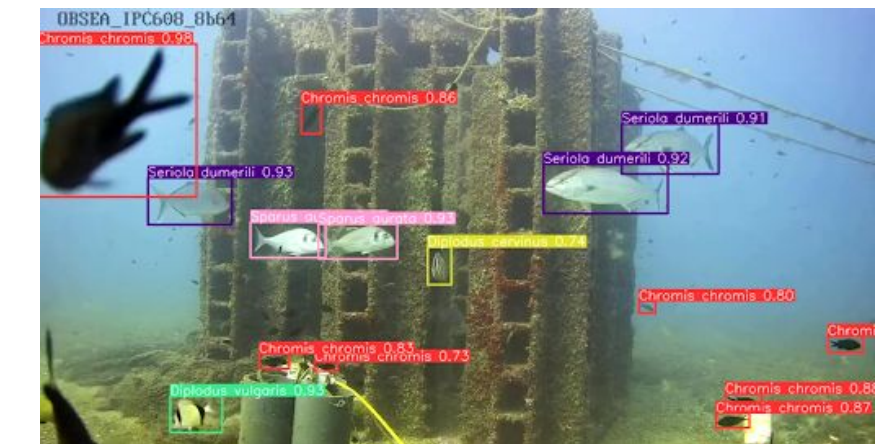


YOLO module

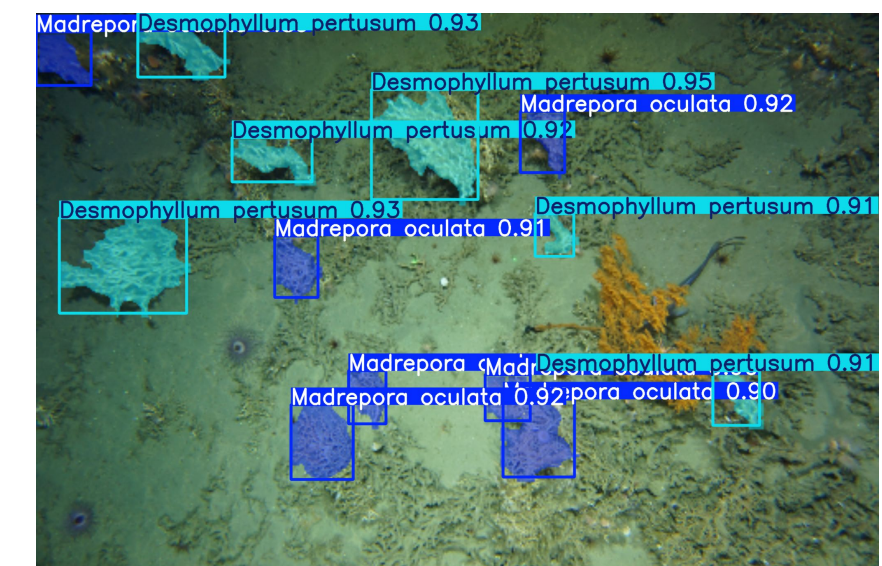
- Based on the Ultralytics' [python package](#). Ultralytics² is a company that develops computer vision models
- YOLO can perform:**
 - Classification:** Identifying what object is in an image
 - Object Detection:** Detecting and localizing multiple objects in an image with bounding boxes
 - Segmentation:** Identifying the exact shape and boundaries of objects in an image (pixel-wise classification)
- YOLO module** on the marketplace:
 - Integrated with [DEEPaaS API](#) (platform support)
 - Dockerized
 - Supports [MLflow](#)
 - Hyperparameter optimization with [Hydra](#), [Optuna](#) & [MLflow](#)



Once retrained
on your data:



OBSEA module able to
detect various fishes



Cold water corals are
properly **segmented**

To-Do: Validation of prototypes

WP3 point-of-view: tech. integration

- To Publish on the marketplace:
 - (very desired) Integrate with DEEPaaS API
 - Create application tests
 - Dockerize
 - Pass CI/CD
- Marketplace makes available:
 - Open source and docker access with rich metadata
 - Try for 10-minutes by everyone
 - Deployable for inference by VO-members
 - Deployable on 3d-party resources via IM

2. Marketplace download
service delivery

1. Marketplace inference
service delivery

3. Inference
service delivery

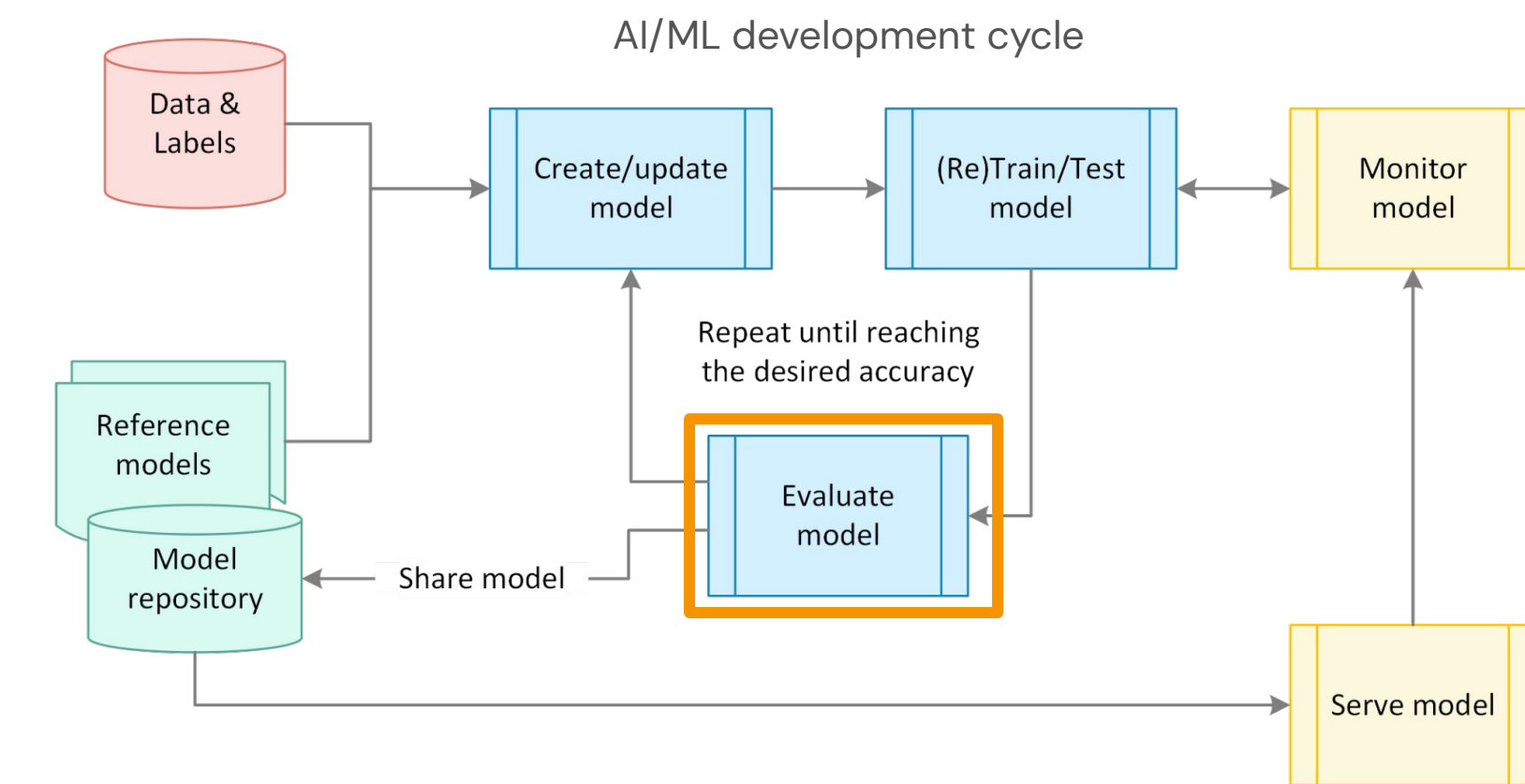
⇒ TRL7 “System prototype demonstration in operational environment”

Metrics: Requested in VA

- a) Number of unique users of the AI image processing service (unit: No. of users)
- b) Number of images processed per year
(= Images uploaded by your end-users for processing by your service)
(unit: No. of images)
- c) Number of images ingested
(= annotated images, made available for download) (unit: No. of images)
- d) Number of countries of users
- e) Names of countries of users

Best Practices:

Evaluate Model



Performance Metrics and Evaluation Methods

- **Accuracy:** correct predictions over total predictions
- **Precision:** ratio of correctly predicted positive observations to the total predicted positives
- **Recall:** ratio of correctly predicted positive observations to all observations in the actual class
- **F1 score:** harmonic mean of precision and recall, useful for imbalanced dataset
- **Intersection over Union (IoU):** overlap between two boxes

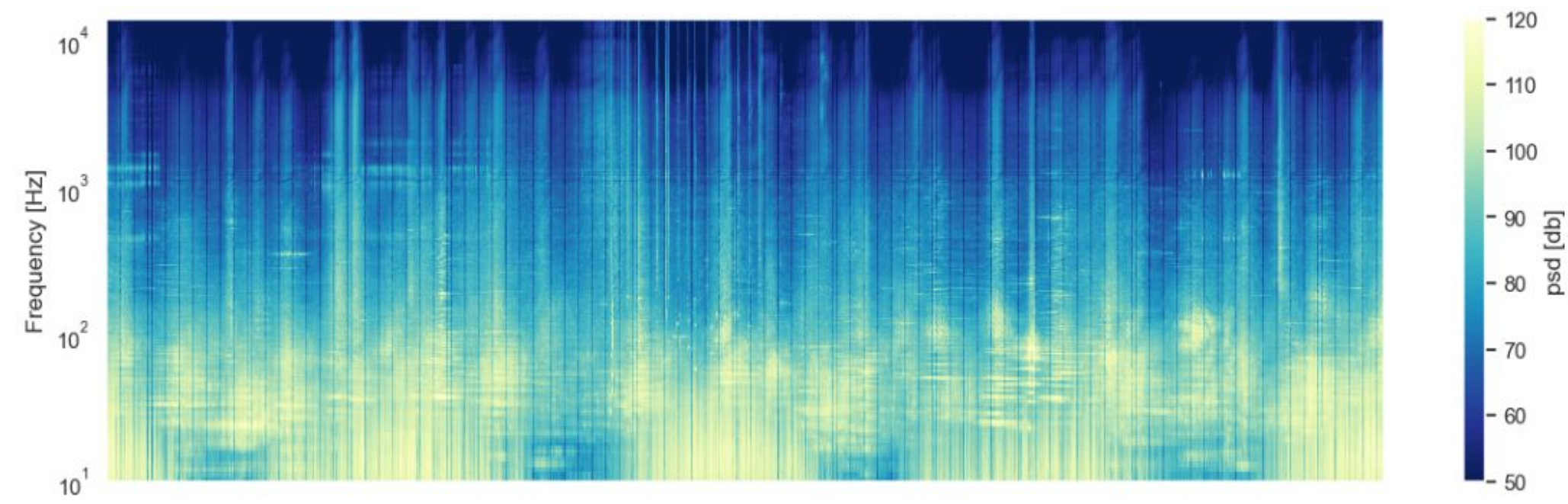
iImagine Use Cases

5 mature + 3 prototype



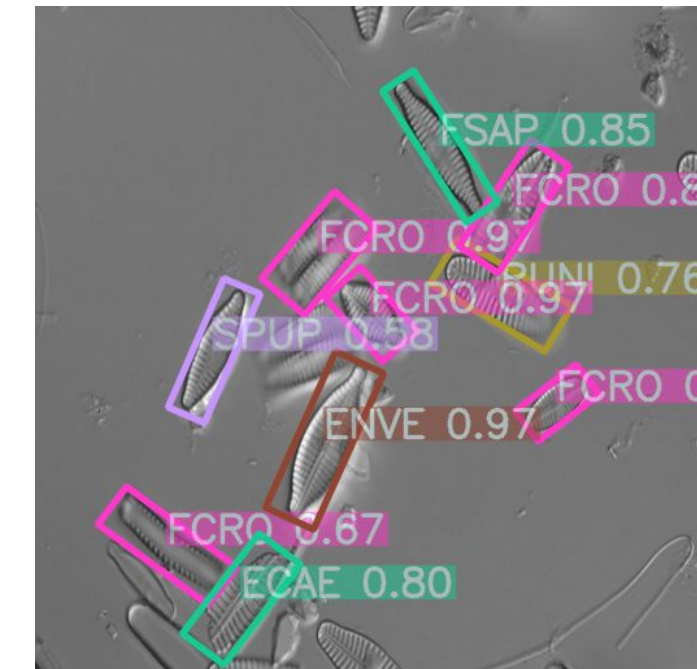
Beach monitoring

SOCIB (ES)



Underwater noise identification

VLIZ (BE)

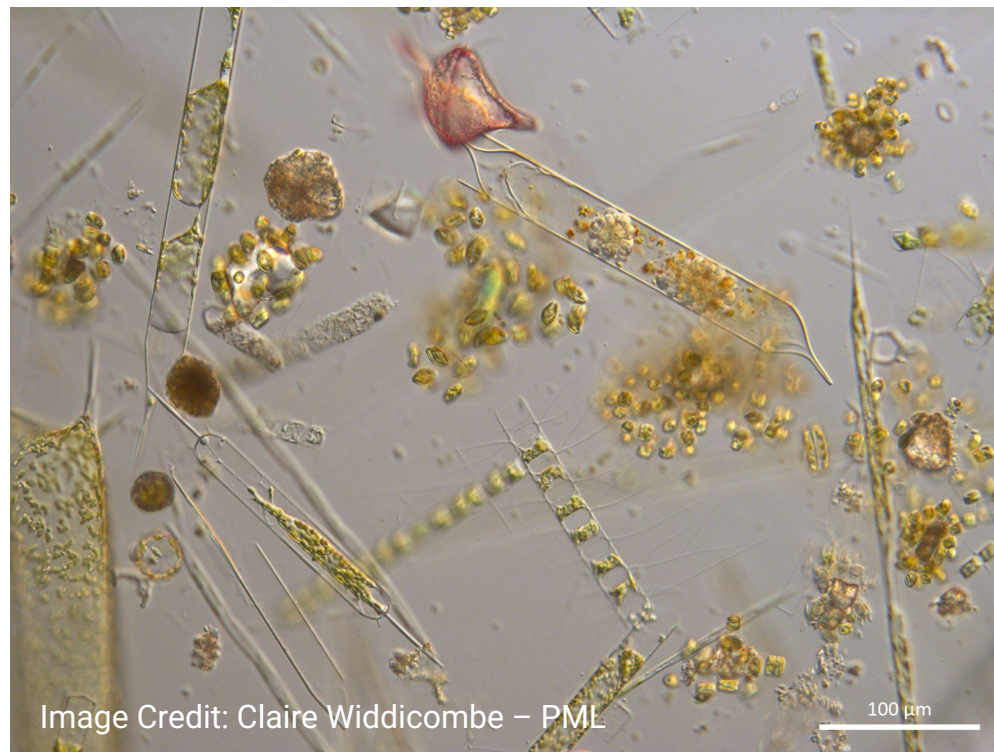


Freshwater diatoms identification

Université de Lorraine-LIEC (FR)

Open Call* Use Cases

Being onboarded via Open Calls



DEcentrAlised Learning (DEAL)

Plymouth Marine Laboratory, UK

Use decentralised learning to increase the robustness of automated plankton and deep-water animal imagery processing



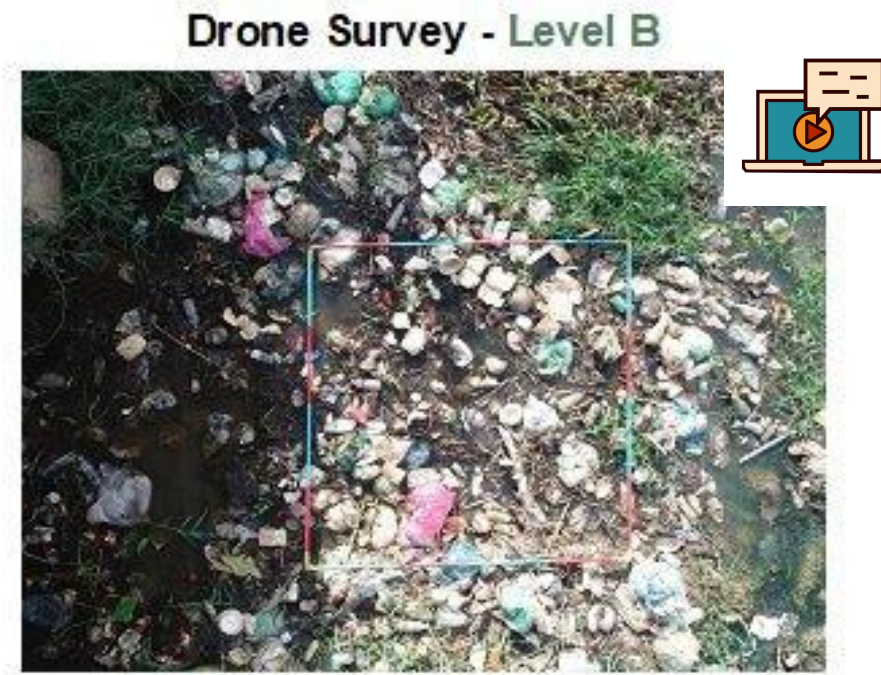
Sea Wave and coastal inundation detection Methodology (SWiM)

LNEC (National Laboratory for Civil Engineering), PT

To enhance a methodology for an automated wave runup and wave detection from camera images for the surveillance of coastline

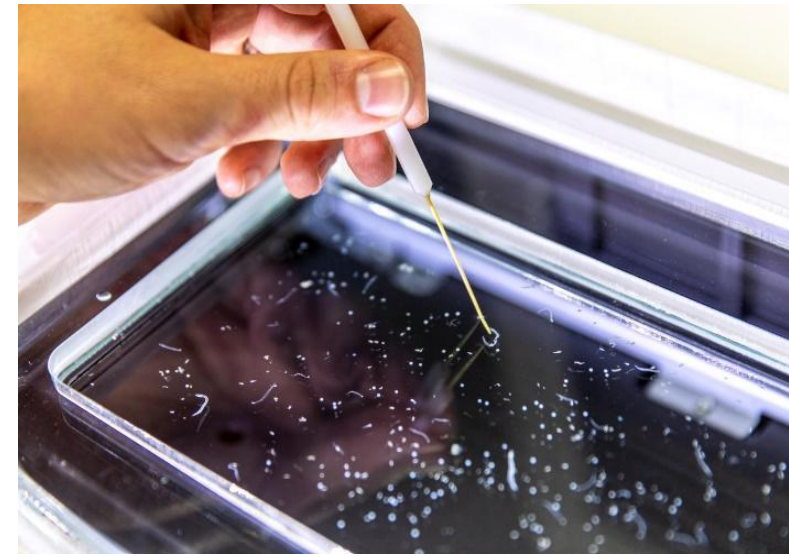
iImagine Use Cases

5 mature + 3 prototype

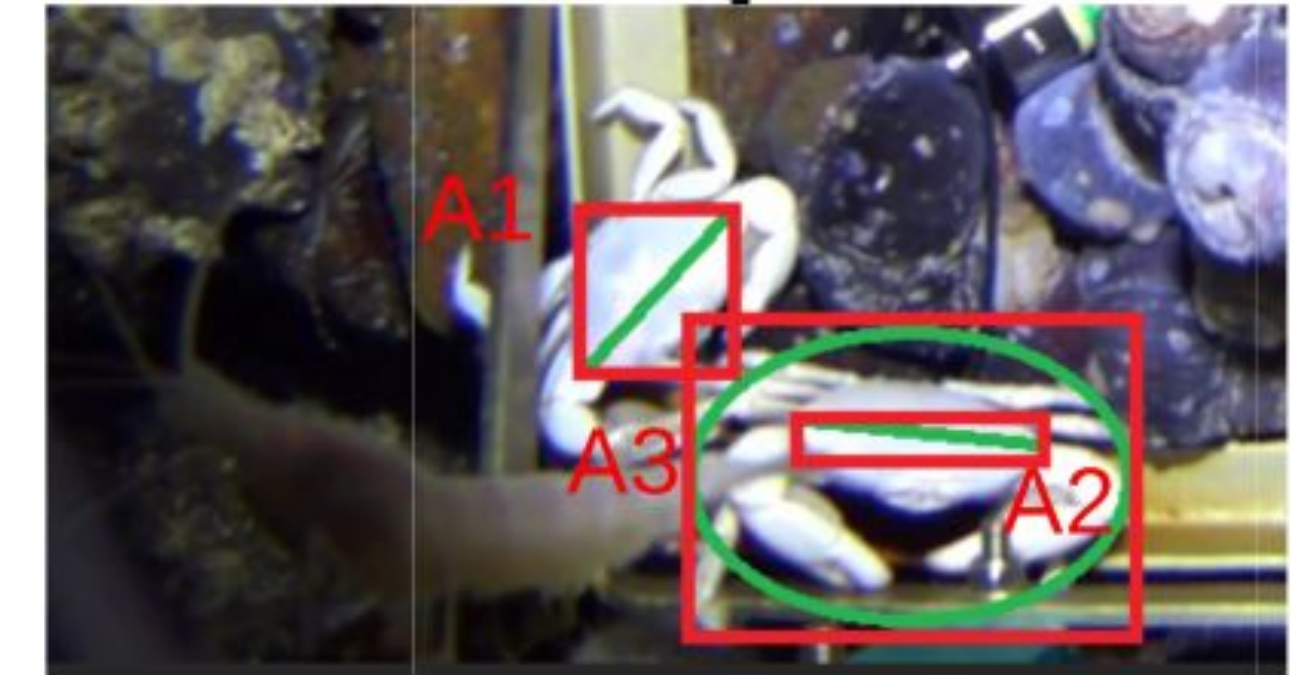


Aquatic Litter Drones

DFKI (DE), MARIS (NL),
OGS (IT)

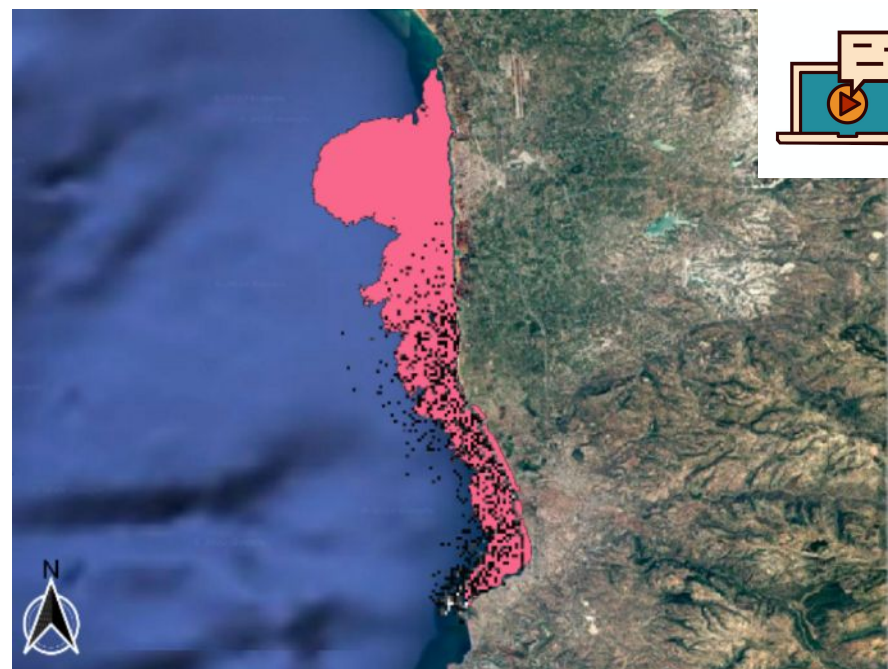


Zooscan – EcoTaxa pipeline (FR)



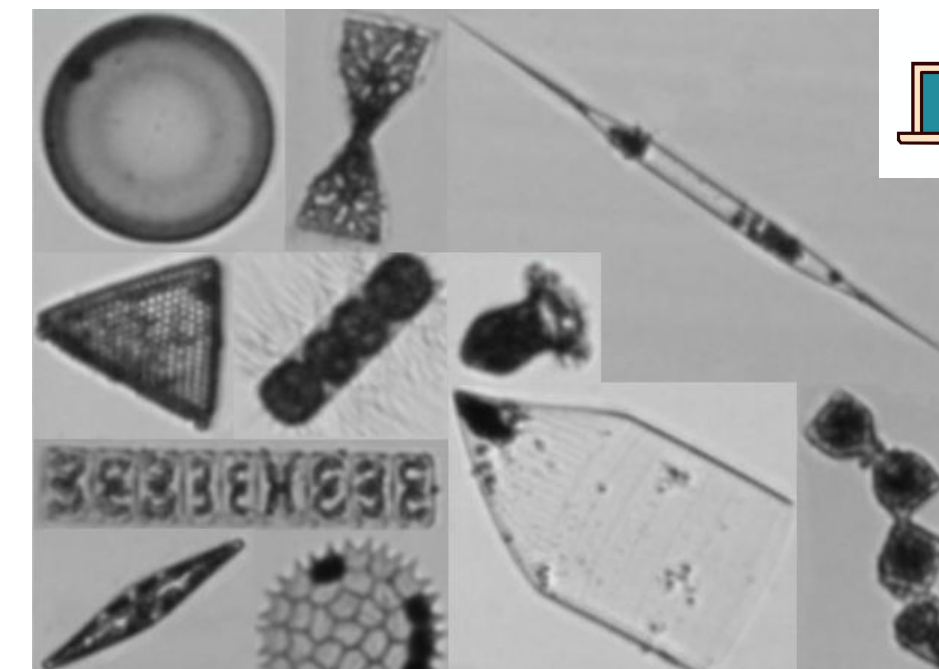
Marine Ecosystem Monitoring

EMSO ERIC:
UPC (ES), IFREMER (FR), MI (IR)



Oil Spill Detection

CMCC, OrbitalEOS, UNITN (IT)

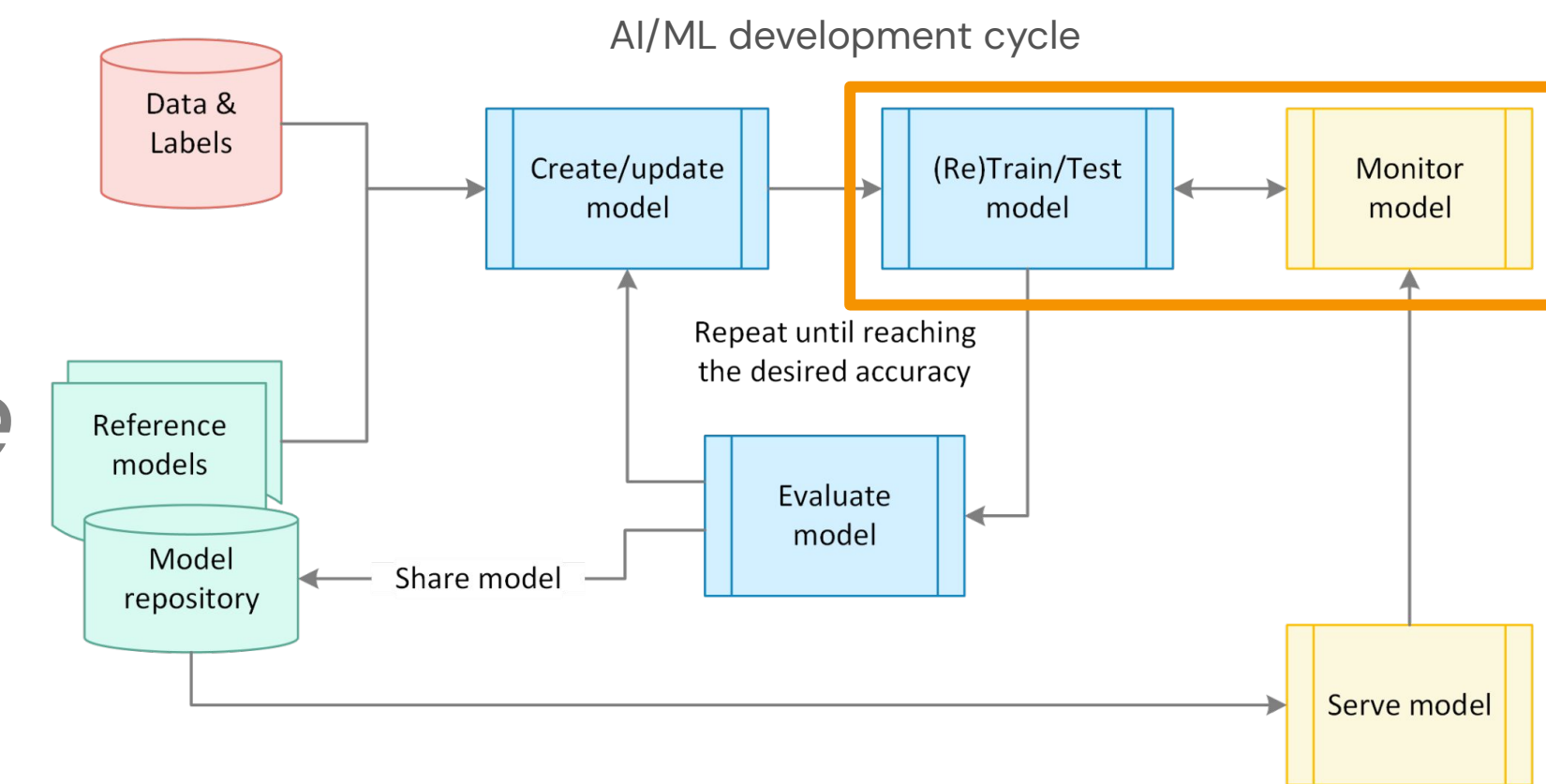


Flowcam phytoplankton identification

VLIZ (BE)

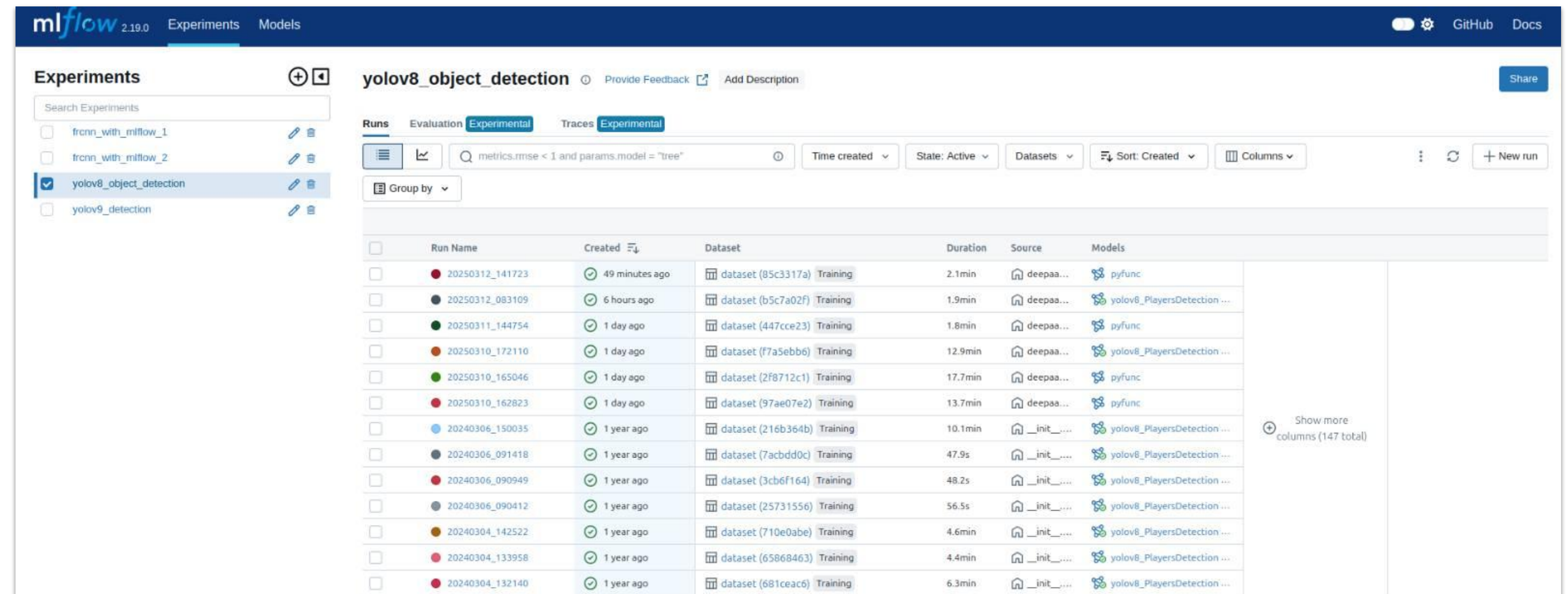
Best Practices:

Monitoring Model Performance



ML Experiment Tracking Tools

- **Tensorboard**: few functionality, single user and self-hosted
- **Weights and Biases**: cloud-based service
- **MLflow**: efficient management and tracking

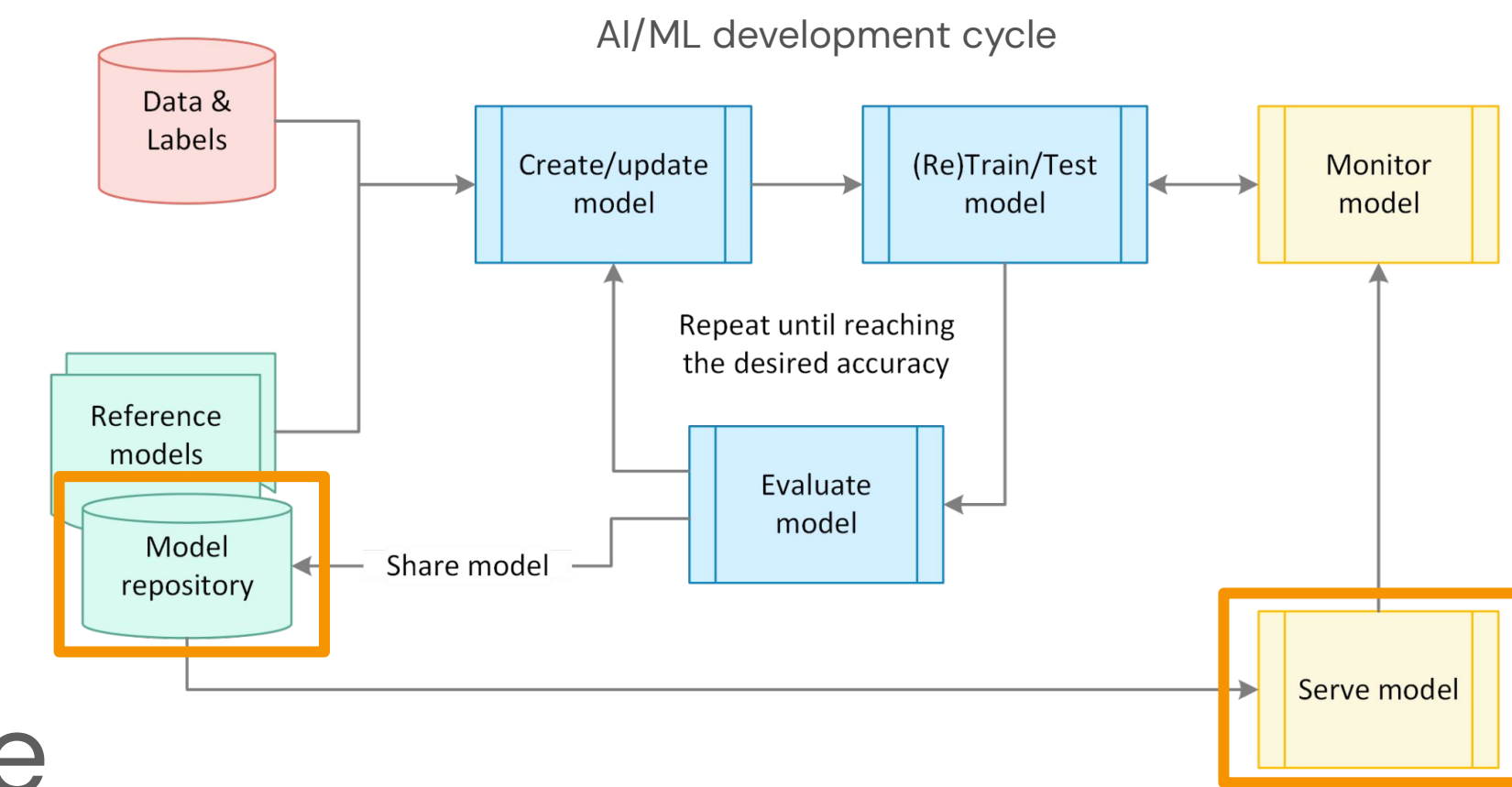


The screenshot shows the MLflow web interface. The left sidebar lists experiments: 'frnn_with_mlflow_1', 'frnn_with_mlflow_2', 'yolov8_object_detection' (selected), and 'yolov9_detection'. The main area displays the 'yolov8_object_detection' experiment details, including a search bar, filters, and a table of runs.

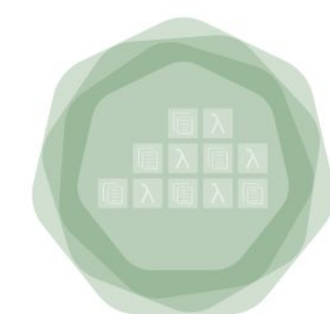
Run Name	Created	Dataset	Duration	Source	Models
20250312_141723	49 minutes ago	dataset (85c3317a) Training	2.1min	deepaa...	pyfunc
20250312_083109	6 hours ago	dataset (b5c7a02f) Training	1.9min	deepaa...	yolov8_PlayersDetection ...
20250311_144754	1 day ago	dataset (447cce23) Training	1.8min	deepaa...	pyfunc
20250310_172110	1 day ago	dataset (f7a5ebb6) Training	12.9min	deepaa...	yolov8_PlayersDetection ...
20250310_165046	1 day ago	dataset (2f8712c1) Training	17.7min	deepaa...	pyfunc
20250310_162823	1 day ago	dataset (97ae07e2) Training	13.7min	deepaa...	pyfunc
20240306_150035	1 year ago	dataset (216b364b) Training	10.1min	_init...	yolov8_PlayersDetection ...
20240306_091418	1 year ago	dataset (7acbdd0c) Training	47.9s	_init...	yolov8_PlayersDetection ...
20240306_090949	1 year ago	dataset (3cb6f164) Training	48.2s	_init...	yolov8_PlayersDetection ...
20240306_090412	1 year ago	dataset (25731556) Training	56.5s	_init...	yolov8_PlayersDetection ...
20240304_142522	1 year ago	dataset (710e0abe) Training	4.6min	_init...	yolov8_PlayersDetection ...
20240304_133958	1 year ago	dataset (65868463) Training	4.4min	_init...	yolov8_PlayersDetection ...
20240304_132140	1 year ago	dataset (681ceac6) Training	6.3min	_init...	yolov8_PlayersDetection ...

Best Practices:

Serve Model



- Sharing and Serving AI Models via iImagine Marketplace
 - AI models are published as Docker images and code repositories
 - Anyone authenticated can **Try** an inference endpoint (for 10 min)
- Model Deployment in Production
 - OSCAR: efficient and scalable deployment for running the AI model inferences



OSCAR

