



Machine Learning Operations (MLOps): from global landscape to practice in AI4EOSC

EGI Conference 2024. October 3d, 2024. Lecce, Italy

<u>Valentin Kozlov^{1),}*</u>, Lisana Berberi¹⁾, Borja Esteban Sanchis¹⁾, Giang Nguyen²⁾, Judith Sainz-Pardo Diaz³⁾, Amanda Calatrava⁴⁾, Germán Molto⁴⁾, Viet Tran²⁾, Alvaro Lopez Garcia³⁾

*<u>valentin.kozlov@kit.edu</u>

1) KIT 2) IISAS 3) IFCA-CSIC 4) UPV



Funded by the European Union

Union 03 | 10 | 2024 by Valentin Kozlov







- Machine Learning Operations (MLOps) definition(s)
- MLOps landscape of platforms & tools
- AI4EOSC MLOps practices









<u>Wikipedia</u>: MLOps is a paradigm that aims to **deploy** and **maintain** machine learning models in **production** reliably and efficiently.

<u>Google</u>: MLOps is an ML engineering **culture** and **practice** that aims at unifying ML system **development** (Dev) and ML system **operation** (Ops).

<u>Databricks</u>*): MLOps is the set of **processes** and **automation** for managing **data**, **code** and **models** to improve performance stability and long-term efficiency in ML systems

*) The Big Book of MLOps: Second Edition (Databricks)







MLOps

MLOps:

culture, practices, processes of automation for managing data, development, models, operations









**** * * ****

MLOps platforms & tools (subset)



7

Al4 meosc

MLOps Platforms

Criteria to assess MLOps Platforms*)

Only **open source** is considered

- Orchestration (0)
- Distributed Training (DT)
- Code Management (CM)
- Model Development (MDV)
- Model Testing/Validation (MTV)

- Model Inference (MI)
- Model Deployment (MDP)
- Experiment Tracking and Metadata Store (ETMS)
- Data Versioning and Management (DVM)
- Model Performance Monitoring (MPM)

Real-world implementation depends on your needs and other already implemented services

*) from L.Berberi "Machine Learning Operations Landscape: Platforms and Tools", submitted



Product	GitHub Stars	O Orche- stration	DT Distri- buted Train- ing	CM Code Man- age- ment	MDV Model Devel- opment	MTV Model Test- ing/- Valida- tion	MI Model Infer- ence	MDP Model Deploy- ment	ETMS Experiment Tracking and Meta- data Store	DVM Data Ver- sioning and Man- agement	MPM Model Per- formance Monitoring	Full Score	Partial Score
MLflow	17.1 K			\checkmark			$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$			30%	10%
Prefect	$14.4~\mathrm{K}$	$\checkmark\checkmark$		\checkmark						$\checkmark\checkmark$		20%	10%
Kubeflow	$13.6~{\rm K}$	$\checkmark\checkmark$	$\checkmark\checkmark$		$\checkmark\checkmark$				\checkmark			40%	0%
Dagster	$10 \mathrm{K}$	$\checkmark\checkmark$						$\checkmark\checkmark$		$\checkmark\checkmark$		30%	0%
W&B (WB)	$8.1~{ m K}$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$		\checkmark	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	70%	10%
MetaFlow	$7.5~{ m K}$	$\checkmark\checkmark$							$\sqrt{}$	\checkmark	\checkmark	20%	20%
Mage	$6.9~\mathrm{K}$						$\checkmark\checkmark$	\checkmark	\checkmark	$\checkmark\checkmark$	\checkmark	20%	30%
Pachyderm	$6.1~{ m K}$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	\checkmark	$\checkmark\checkmark$		60%	30%
ClearML	$5.2~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	100%	0%
Flyte	$4.7~\mathrm{K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$		80%	10%
Seldon core	$4.2~\mathrm{K}$	$\checkmark\checkmark$				$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark		\checkmark	50%	10%
ZenML	$3.6~{\rm K}$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	\checkmark	100%	0%
Polyaxon	$3.5~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	\checkmark	\checkmark	90%	10%
TFX	$2.1~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$		70%	20%
MLeap	$1.5~{ m K}$	$\checkmark\checkmark$						$\checkmark\checkmark$			$\checkmark\checkmark$	30%	0%
MLRun	$1.2 \mathrm{~K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	80%	20%

Product	GitHub Stars	O Orche- stration	DT Distri- buted Train- ing	CM Code Man- age- ment	MDV Model Devel- opment	MTV Model Test- ing/- Valida- tion	MI Model Infer- ence	MDP Model Deploy- ment	ETMS Experiment Tracking and Meta- data Store	DVM Data Ver- sioning and Man- agement	MPM Model Per- formance Monitoring	Full Score	Partial Score	
MLflow	17.1 K			\checkmark			$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark			30%	10%	
Prefect	$14.4~\mathrm{K}$	$\checkmark\checkmark$		\checkmark						$\checkmark\checkmark$		20%	10%	
Kubeflow	$13.6~{\rm K}$	$\checkmark\checkmark$	$\checkmark\checkmark$		$\checkmark\checkmark$				\checkmark			40%	0%	
Dagster	$10 \mathrm{K}$	$\checkmark\checkmark$						$\checkmark\checkmark$		$\checkmark\checkmark$		30%	0%	
W&B (WB)	$8.1~{ m K}$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$		\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	70%	10%	
MetaFlow	$7.5~{ m K}$	$\checkmark\checkmark$							\checkmark	\checkmark	\checkmark	20%	20%	
Mage	$6.9~{ m K}$						$\checkmark\checkmark$	\checkmark	\checkmark	$\checkmark\checkmark$	\checkmark	20%	30%	
Pachyderm	$6.1~{ m K}$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	\checkmark	$\checkmark\checkmark$		60%	30%	
ClearML	$5.2~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	100%	0%	
Flyte	$4.7~\mathrm{K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$		80%	10%	
Seldon core	$4.2~\mathrm{K}$	$\checkmark\checkmark$				$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark		$\checkmark\checkmark$	50%	10%	
ZenML	$3.6~{\rm K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	100%	0%	
Polyaxon	$3.5~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	\checkmark	$\checkmark\checkmark$	90%	10%	
TFX	$2.1~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$		70%	20%	
MLeap	$1.5~{ m K}$	$\checkmark\checkmark$						$\checkmark\checkmark$			$\checkmark\checkmark$	30%	0%	
MLRun	$1.2~{\rm K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	80%	20%	
very limited o	open-sourc	e version											1	0

Product	GitHub Stars	O Orche- stration	DT Distri- buted Train- ing	CM Code Man- age- ment	MDV Model Devel- opment	MTV Model Test- ing/- Valida- tion	MI Model Infer- ence	MDP Model Deploy- ment	ETMS Experiment Tracking and Meta- data Store	DVM Data Ver- sioning and Man- agement	MPM Model Per- formance Monitoring	Full Score	Partial Score	
MLflow	17.1 K			\checkmark			$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark			30%	10%	
Prefect	$14.4~\mathrm{K}$	$\checkmark\checkmark$		\checkmark						$\checkmark\checkmark$		20%	10%	
Kubeflow	$13.6~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$		$\checkmark\checkmark$				\checkmark			40%	0%	
Dagster	$10 \mathrm{K}$	$\checkmark\checkmark$						$\checkmark\checkmark$		$\checkmark\checkmark$		30%	0%	
W&B (WB)	$8.1~{ m K}$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$		\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	70%	10%	
MetaFlow	$7.5~{ m K}$	$\checkmark\checkmark$							\checkmark	\checkmark	\checkmark	20%	20%	
Mage	$6.9~{ m K}$						$\checkmark\checkmark$	\checkmark	\checkmark	$\checkmark\checkmark$	\checkmark	20%	30%	
Pachyderm	$6.1~{ m K}$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	\checkmark	$\checkmark\checkmark$		60%	30%	
ClearML	$5.2~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	100%	0%	
Flyte	$4.7~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$		80%	10%	
Seldon core	$4.2~{ m K}$	$\checkmark\checkmark$				$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark		$\checkmark\checkmark$	50%	10%	
ZenML	$3.6~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	100%	0%	
Polyaxon	$3.5~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	\checkmark	$\checkmark\checkmark$	90%	10%	
TFX	$2.1~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$		70%	20%	
MLeap	$1.5~{ m K}$	$\checkmark\checkmark$						$\checkmark\checkmark$			$\checkmark\checkmark$	30%	0%	
MLRun	$1.2~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	80%	20%	
very limited of	open-sourc	ce version		K8s bc	ound								1	1
from L.Berbe	eri "Machi	ne Learni	ng Opera	tions Lan	idscape: F	Platforms	and Tools	s", submitt	ed	///				11

Т

Product	GitHub Stars	O Orche- stration	DT Distri- buted Train- ing	CM Code Man- age- ment	MDV Model Devel- opment	MTV Model Test- ing/- Valida- tion	MI Model Infer- ence	MDP Model Deploy- ment	ETMS Experiment Tracking and Meta- data Store	DVM Data Ver- sioning and Man- agement	MPM Model Per- formance Monitoring	Full Score	Partial Score	
MLflow	17.1 K			\checkmark			$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark			30%	10%	
Prefect	$14.4~\mathrm{K}$	$\checkmark\checkmark$		\checkmark						$\checkmark\checkmark$		20%	10%	
Kubeflow	$13.6~{\rm K}$	$\checkmark\checkmark$	$\checkmark\checkmark$		$\checkmark\checkmark$				\checkmark			40%	0%	
Dagster	$10 \mathrm{K}$	$\checkmark\checkmark$						$\checkmark\checkmark$		$\checkmark\checkmark$		30%	0%	
W&B (WB)	$8.1~{ m K}$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$		\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	70%	10%	
MetaFlow	$7.5~{ m K}$	$\checkmark\checkmark$							\checkmark	\checkmark	\checkmark	20%	20%	
Mage	$6.9~{ m K}$						$\checkmark\checkmark$	\checkmark	\checkmark	$\checkmark\checkmark$	\checkmark	20%	30%	
Pachyderm	$6.1~{ m K}$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	\checkmark	$\checkmark\checkmark$		60%	30%	
ClearML	$5.2~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	100%	0%	
Flyte	$4.7~\mathrm{K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$		80%	10%	
Seldon core	$4.2~\mathrm{K}$	$\checkmark\checkmark$				$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark		$\checkmark\checkmark$	50%	10%	
ZenML	$3.6~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	100%	0%	
Polyaxon	$3.5~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	\checkmark	$\checkmark\checkmark$	90%	10%	
TFX	$2.1~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$		70%	20%	
MLeap	$1.5~{ m K}$	$\checkmark\checkmark$						\checkmark			$\checkmark\checkmark$	30%	0%	
MLRun	$1.2~{ m K}$	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	\checkmark	$\checkmark\checkmark$	$\checkmark\checkmark$	80%	20%	
very limited o	open-sourc	e version		K8s bc	und	TF-bo	ound	Builds	on top of othe	er tools			1	0
from L.Berbe	eri "Machii	ne Learni	ng Opera	tions Lan	idscape: F	Platforms	and Tools	s", submitt	ed					Z



MLOps Drift tools

Criteria to assess Drift tools*)

Only **open source** is considered.

Types of drift detectors reviewed:

- Change drift : monitor single variables in the streaming context
- Concept drift : monitor the performance of the model, trying to identify shifts in the learned concept, i.e. between the data's feature values and their labels
- Data drift : monitor the distribution of the data features (model-agnostic)
- *Ensemble*: groups of detectors combined to draw the conclusion

Two data modes are reviewed: *streaming* and *batch*

*) from L.Berberi "Machine Learning Operations Landscape: Platforms and Tools", submitted



			River		Alibi Detect		Torchdrift		t Frouros		Men	elaus	Evic	lently
$egin{array}{c} \mathbf{Drift} \ \mathbf{Type} \end{array}$	Drift Detector Algorithm	Abbrev. Ref.	${{\it Stream}}_{ing}$	Batch	Stream-ing	Batch	Stream- ing	Batch	Stream- ing	Batch	Stream- ing	Batch	Stream- ing	Batch
nge ction	Bayesian Online Change Detection Cumulative Sum Test Geometric Moving Average	BOCD, (Adams and MacKay, 2007) CUSUM, (Page, 1954) - (Roberts, 1959)	\$ \$ \$		√ √				√ √ √		V		~	
Cha Dete	Detection Page-Hinkley ADaptive WINdowing	PH, (Page, 1954) ADWIN, (Bifet and Gavaldà, 2007)	v v						√ √		\checkmark		\checkmark	
	Kolmogorov-Smirnov Win- dowing Detection	KSWIN, (Raab et al, 2020)	\checkmark		\checkmark		\checkmark		\checkmark					\checkmark
	Drift Detection Method Early Drift Detection Method	DDM, (Gama et al, 2004) EDDM, (Baena-Garcia et al, 2006)	√ √						√ √		√ √			
t Drift	EWMA Concept Drift Detec- tion Warning Statistical Test of Equal Pro- portions to Detect concept	ECDDWT, (Ross et al, 2012) STEPD, (Nishida and	\checkmark						√ √		√			
Concept	drift Hoeffding's drift detection Method Fast Hoeffding drift detec- tion	Yamauchi, 2007) HDDM, (Frías-Blanco et al, 2015) FDDM, (Pesaranghader and Viltor 2016)	\checkmark						√ √					
	Reactive Drift detection Method	RDDM, (Barros et al, 2017)							\checkmark					
	Cramér-von Mises test Method	CVMTest, (Cramér, 1928)		\checkmark		\checkmark				\checkmark		\checkmark		\checkmark
IJ	Hellinger Distance Drift Detection Method Kullback-Leibler divergence	HIDDDM, (Hdlinger, 1909) KL, (Kullback and Leibler,		\checkmark		/				√	/	٠ •		√
ta Dri	Detection PCA-Based Change Detec- tion	1951) PCA-CD, (Cramér, 1928)	\checkmark			✓ ✓		\checkmark		v	√ √	V		√ √
Da	Earth Mover's Distance Detection Method	EMD, (Rubner et al, 2000)								\checkmark				\checkmark
	Maximum Mean Discrepancy Detection	MMD, (Gretton et al, 2012)	\checkmark		\checkmark		\checkmark		\checkmark	\checkmark				\checkmark
	Incremental Kolmogorov- Smirnov	IncrementalKSTest, (dos Reis et al, 2016)	\checkmark		\checkmark				\checkmark					\checkmark
n- nble	Streaming Ensemble	-, (Maciel et al, 2015)	\checkmark		\checkmark				\checkmark		\checkmark		\checkmark	
En	Batch Ensemble	-		\checkmark		\checkmark						\checkmark		\checkmark
		Gitsource	https:// .com/c l/r	//github online-m river	http .cor O/a	s://github n/SeldonI libi-detect	https:/ .com/t t/tor	//github corchdrif chdrift	https:// .com/l dvance puting	//github IFCA-A ed-Com /frouros	https:// .com/r ene	/github nitre/m elaus	https:/ .com/e ai/ev	//github evidently ridently

2 . 10			Ri	ver	Alibi Detect		Torc	hdrift	Fro	uros	Men	elaus	\mathbf{Evid}	lently
$egin{array}{c} { m Drift} \ { m Type} \end{array}$	Drift Detector Algorithm	Abbrev. Ref.	${{\it Stream}}_{ing}$	Batch	${{ Stream} \atop {ing}}$	Batch	Stream- ing	Batch	Stream- ing	Batch	Stream- ing	Batch	Stream- ing	Batch
Change Detection	Bayesian Online Change Detection Cumulative Sum Test Geometric Moving Average Detection Page-Hinkley ADaptive WINdowing Kolmogorov-Smirnov Win- dowing Detection	BOCD, (Adams and MacKay, 2007) CUSUM, (Page, 1954) -, (Roberts, 1959) PH, (Page, 1954) ADWIN, (Bifet and Gavaldà, 2007) KSWIN, (Raab et al, 2020)			√ √ √		v				√ √ √		۷ ۷	✓
Concept Drift	Drift Detection Method Early Drift Detection Method EWMA Concept Drift Detec- tion Warning Statistical Test of Equal Pro- portions to Detect concept drift Hoeffding's drift detection Method Fast Hoeffding drift detec- tion Reactive Drift detection Method	DDM, (Gama et al, 2004) EDDM, (Baena-Garcia et al, 2006) ECDDWT, (Ross et al, 2012) STEPD, (Nishida and Yamauchi, 2007) HDDM, (Frías-Blanco et al, 2015) FDDM, (Pesaranghader and Viktor, 2016) RDDM, (Barros et al, 2017)	√ √ √								√ √ √			
En- semble	Cramér-von Mises test Method Hellinger Distance Drift Detection Method Kullback-Leibler divergence Detection PCA-Based Change Detec- tion Earth Mover's Distance Detection Method Maximum Mean Discrepancy Detection Incremental Kolmogorov- Smirnov Streaming Ensemble Batch Ensemble	CVMTest, (Cramér, 1928) HIDDDM, (Hdlinger, 1909) KL, (Kullback and Leibler, 1951) PCA-CD, (Cramér, 1928) EMD, (Rubner et al, 2000) MMD, (Gretton et al, 2012) IncrementalKSTest, (dos Reis et al, 2016) -, (Maciel et al, 2015)	√ √ √	√ √ √	√ √ √	√ √ √	~	V	✓ ✓ ✓		√ √ √	√ √ √	√	
		Gitsource	https:/ .com/c l/r	//github online-m iver	http .cor O/a	s://github n/SeldonI libi-detect	https:/ .com/t t/tor	/github orchdrif chdrift	https:/ .com/l dvance puting	/github FCA-A ed-Com /frouros	https:/ .com/r ene	/github nitre/m elaus	https:/ .com/e ai/ev	/github vidently idently



AI4EOSC practices







AI4EOSC practices



Al4 meosc

AI4EOSC practices, MLflow

- Multi-user instances for <u>AI4EOSC</u> and <u>iMagine</u>
- Self-registrations service based on VO membership
 - Access update
 - Experiment & Model permissions
- MLflow's Support for LLMs (next talk by L.Berberi)
- Suitable for Federated Learning Experiment Tracking (after next <u>talk by K.Alibabaei</u>)
- Usage monitoring with Goaccess
- Regular Backups & Garbage collection



Funded by the European Union

MLFlow	Regi	stra	tion			WILFI	owu	ser	Settin	igs
Email: valentin.kozlov Registered: No	v@gmail.c	com				UserID: 6 Email: valer Registered: Admin: Yes	tin.kozlov(Yes	@kit.edu		
Outonto						Password:				
Create accou	Int							Please ch	ioose a secure j	basswor
Password:									U	pdate
I accept the Terms of	o <mark>f Use</mark> and Please cho	l <u>Privacy</u> ose a seci	r Policy			Experime	nt perm	issions		
			Create			Experiment I	0	User ID	Permissio	n
						9		6 (You)	Manage	
	Logout					14		6 (You)	READ	
Privacy Po	olicy Term	s of Use				15		6 (You)	READ	
Search Experiments Contralized_Learning MrTare_Sinulation_HoreKa MrTare_Sinulation_HoreKa MrTare_HoreKa_Halcove MrTare_Flowe_Text		Experiment C Desc Thermal Q metri Fig. Sort. 6 Table C	ID: 83 Artifact Location: million liption Ed: urban feature semantic seg comme < 1 and personamodel = " intented " [] Okuma >] part Evaluation Experimentation	v-artifacts;k gmentatic 'tree'	on ()	Time created *	te: Active v	Datasets 👻	Ŧ	0 0
Search Experiments Centralized_caving NMTexe_Simulation_transfe NMTexe		Experiment	Du B3 Artfact Location: mittion iption Eat: urban feature semantic seg- control of the semantic seg- control of the semantic seg- control of the semantic seg- control of the semantic seg- tion takes Bio Name Set-Securico Li Set-Securico Li Set-Securi	smentatic gmentatic treat O 1 O 1 O 1 O 1 Create	ed Right States	Time created v Sta Dataset - -	te: Active v	Datasets v Daration 1.6h 1.6h 1.2h	E Source	C E Models S tensorfi S tensorfi S tensorfi
Secto Experiments Committee Constraints Notine, Simulation, Learning Notine, Simulation, Learning Notine, Constraints, Simulation, Learning Notine, Simulation, L		Experiment Desc Thermal C matrix	Do 13 Artifact Location mittion (giption dat urban feature semantic age urban feature semantic age urban feature semantic age Do teace Do teace Sector (Sector)(D) Setor (Sector)(D) Setor)(D) Setor)(D) Setor (Sec	v-artifacts, k	ed 🖗 morth ago	Time created v Sta Dataset - - - -	te: Active v	Durasets V Duratism 1.6h 1.2h 1.2h 1.2h	i Source - - - -	Models B tensorff B tensorff B tensorff
Search Departments Campada Lannia Withing, Smaller U. Horiz Withing, Young, Small Withing, Young, Small Withing, Young, Small		Experiment V Desc Thermal Q matri \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$ \$	0.0.13 Arthac Location mitton iption rat: untran features semantic seg untran < 1 and parent model untran < 1 and parent model untra < 1 and parent model untra < 1 and parent model untra < 1 and parent model Nan Tetre Para Tetre Para Tetre Para ScoreGO 0.0000 - 1000000 - 100000 0.0000 - 000000 - 100000 0.0000 - 00000 - 100000 0.0000 - 00000 0.0000 - 00000 0.00000 0.0000 - 00000 0.000000 0.000000 0.00000 0.000000 0.000000 0.000000 0.000000 0.000000 0.00000000	smentatic	ed Fig. meth ago menth ago menth ago menth ago menth ago	Time created v Sta Dataset - - - - -	te: Active v	Duration Duration 1.6h 1.2h 1.2h 40.6min 40.6min 57.5min 57.5min	5 5 6 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7 7	Models Stonsorf Stonsorf Stonsorf - - -
Sean Daymons Caracterization Withing Smaller Union Withing Smaller Union Withing Smaller Union Withing Smaller Union Withing Smaller Union Withing Smaller Union Withing Smaller Union		Experiment Construction Cons	0.10.3 Artika Loszlów mitłow ipistim sat: urban festurow semantic seg urban - 1 und garene mitłoł urban - 1 und garene mitłoł u	Create Cr	ed 74 emerch ago emerc	Time created v Sa Dataset - - - - - - - - -	te: Active v	Duration Duration 1.6h 1.2h 1.2h 1.2h 1.2h 1.2h 1.2h 1.2h 1.2	* Ssurce - - - - - - -	Models Bitosart Bitos
Sean Experiments Control Control Control On Control Control On Con	/ 0 / 0 / 0 / 0	Experiment	D. 13 Artise Loadiour with a pipton issue that the second of the second of the second of the second of the example of the second of the example of the e	amentatic see" Create 0 11 0 11 0 11 0 11 0 11 1 0 11 1 0 11 1 0 11 0 1	ed Pa merch ago merch ago merch ago merch ago merch ago merch ago merch ago	Time or seted v 30	te: Active v)	Duration 1.6n 1.6n 1.2n 1.2n 1.2n 40.6nn 40.6nn 40.6nn 87.5nn	Searce	Models B tonsart C C C C C C C C C C C C C C C C C C C
Event Developments Grademanning Grademanning Grademanning Grademanning Writes_Transform Writes_Transform Writes_Transform Ordeman_transform Ordeman_transform Ordeman_transform Stateman_transform Stat	A Construction of the second s	Experiment C matri Fig Sort C C C C C C C C C C C C C	D. D. A Artige Leading with grate faith that seture semantic equipation for any setup and any setup hold of any setup any setup hold of any setup any setup hold of any se	gmentatic see " Crease 0 1s 0 1s 0 1s 0 1s 0 1s 0 1s 0 1s 0 1s	ed Fig.	The crush v for	te: Active v)	Duratism Duratism 1.6n 1.2h 1.2h 1.2h 1.2h 1.2h 1.2h 1.2h 1.2h	: Searce - - - - - - - - - - - - - - - - - - -	Models B tensor B tensor C Services C
Send Experiment Grade Section Grade	S Muser for several States of the several States of the several States of the several States of the several States of the several States of the several States of the several St	Eperiment	D. 13 Artifac Loadian within pipton issue The Network semantics equilibrium Internet and annum Artifaction Internet Intern	gmentatic tree* Create 0 11 0 11 0 11 0 11 0 11	ed F2 method F2 method F2 method F3 method F3 meth	Time crusted w 50	te: Active v)	Duratism 1.6h 1.6h 1.2h 1.2h 1.2h 1.2h 1.2h 1.2h 40.6m 1.2h 40.6m 1.2h 40.6m 1.2h	I Searce 	Models Tosself Toss

18

Al4 meose

Full project name: AI40S-hub/litter-assessment/main

AI4EOSC practices, CI/CD

CI/CD and inference

Stage View

	Declarative: Checkout SCM	Metadata tests	AI4OS Hub metadata V1 validation	AI4OS Hub metadata V2 validation (JSON)	AI4OS Hub metadata V2 validation (YAML)	License validation	Check if only metadata files have changed	User- defined module pipeline job	Docker build and delivery	Docker Variable initialization	AI4OS Hub Docker images build	AI4OS Hub Docker delivery to registry	Update OSCAR services	Declarative: Post Actions
Average stage times: (Average <u>full</u> run time: ~16min	2s	744ms	20s	0ms	24s	2s	1s	7min 14s	368ms	2s	5min 51s	2min 51s	14s	691ms
#20 56s) Sep 23 8 15:36 commits	2s	744ms	20s		24s	2s	1s	7min 14s	368ms	2s	5min 51s	2min 51s	14s	691ms

		OSCAR		≡ <		Valentin Kozlov - https://i	/alentin Kozlov - https://inference-walton.cloud.imagine-ai.eu []				
	:: •	Services Minio Storage body-pose-test	^	Bucket litter-asses	ssment:		Search			م	
		dogbreed		SELECT FILES PUP	PLOAD 🔥	¢					_
Funded by		grayify-public				↑ Name	Size		Last Modified	~ ~	
the European Union		lit litter-assessment		÷		in out				Q 0	19
											/

Al4 OCC

AI4EOSC practices, Drift monitoring

(Work-in-Progress)

- Release of a new dataset:
 => Before publishing, asses for the drift (e.g. DVC + CI/CD)
- New data at the production service:
 => in parallel with inference, run a drift detector

e.g. J. Sisniega et al, Fut.Gen.Comp.Sys.,161:174-188, 12 2024

It is useful to have historical overview of drift detector runs with relevant metadata









- MLOps is a modern practice of automation for managing data, development, models, operations of AI/ML/DL-based services
- More MLOps tools and platforms become available, choose those you really need
- AI4EOSC services already cover a good part of MLOps processes, Level 1











Co-funded by the European Union



Reach us!

Thank you for your attention Project Coordinator: Álvaro López García - aloga@ifca.unican.es

