Bratislava Workshop – 15/11/2023

AI4 MOSC

AI4Compose: low-code composition of AI inference pipelines

Amanda Calatrava Universitat Politècnica de València (UPV)



Objectives



Feature rich services and platform to build and deploy custom AI applications in the EOSC

Objective 2

Support for building AI systems on distributed datasets, with a particular focus on federated learning

Objective 3

Services to compose AI tools, enabling the development of complex data-driven applications



AI exchange/hub in the context of the EOSC, enhancing and increasing the application offer currently available

AI4

CO-funded by the European Union

Objective 5

Extend the service offer and the capabilities being offered through the EOSC portal, with focus on AI

Goal

Foster an AI exchange in the EOSC context transforming the development of AI applications in the EOSC



Graphical Composition of **AI Inference Pipelines**

AI4EOSC Platform architecture



OSCAR: an overview

AI4 COEOSC Co-funded by the European Union

OSCAR is an open-source platform that:

- Support the serverless (event-driven + automated elasticity) computing model for data-processing applications on Kubernetes.
- Is easily deployed on multi-Clouds via the Infrastructure Manager (IM).
- Supports the execution of highly-parallel event-driven data-processing serverless applications.
- Integrated with IM Dashboard, EGI DataHub, EGI Check-In and dCache (via WebDav).

OSCAR is used for:

• Scalable AI inference of pre-trained AI models.

OSCAR instance for AI4EOSC: <u>https://inference.cloud.ai4eosc.eu#/login</u>



What is Node-RED?

- Node-RED is a flow-based programming tool for connecting hardware devices, APIs, and services.
- Built on NodeJS and the D3.js library.
- The minimal structure are the nodes.
- Nodes are organized in flows that connect nodes.



Al4 meosc

Co-funded by

the European Union

https://nodered.org

What is Elyra?

AI4 COEOSC Co-funded by the European Union

- Elyra is an open-source project for developing and running machine learning workflows in JupyterLab.
- Provides tools for users to create visual pipelines for machine learning workflows.
- The minimum structure for a workflow are the nodes (Python script, R script, and Jupyter notebook).
- Elyra is programmed in Python and allows the use of many libraries focused on data analysis.



https://elyra.readthedocs.io/en/stable/

Al4Compose: Low-code composition of Al inference pipelines.



Al4 meosc

Co-funded by the European Union

https://github.com/ai4os/ai4-compose/tree/main





Al4 meosc

Co-funded by

the European Union

Composing Al pipelines: an example



Rode-RED											-/"	Deploy 🔻	≗ ≡
filter nodes	Deploy Services	Workflows OSCAR	Workflows OSCAR	Workflows 2 invoke	Flow 2	Workflows IA	Example 2 IBERGR	Example 1 IBERGR	Flow 7	+ •	🕸 debug	i 🖉	÷ ¢
subflows												T selected nodes	r 📋 all 🔻
OSCAR Grayity Services (2) OSCAR Plants Services (2)													
common													
inject inject idebug icomplete catch istatus inik call inik call comment													
function													
f tunction switch X change ij range				റ്റത്ത	MAGE / 4	N ന്നരർമ	<u>ചിര നാര</u> ാ	វារិ					
e { template				Sound	hnee t		aid i Gou						
* *	Q									0 +			Ģ

Composing AI pipelines: an example





Al4 Spece Co-funded by the European Union

- New modules to interact with OSCAR, published in the official Node-RED library (more coming soon):
 - <u>https://flows.nodered.org/search?term=AI4EOSC</u>
- <u>Flowfuse</u>: to manage users and multiple instances of Node-RED. Official instance of the project:
 - <u>https://forge.flows.dev.ai4eosc.eu</u>
 - OSCAR templates ready to offer a pre-configured Node-RED instance with OSCAR support.

ode-RED home about blog docur	mentation forum flows gith	uub	FlowFuse			
node-red-contrib-oscar-grayify-service 1.0.0	Node Info Version: 1.0.0		Al4 COEOSC	Indance - Owniewe - Rowniewe × ← ⇒ C is flows:devealeeose FlowFuse	■ Readroge: Reactions: R + execting the second secon	di de
This node supports highly-scalable event-driven image conversion to grayscale using the popular ImageMagick software.	Updated 1 week ago License: ISC Rating: <i>not yet rated</i> View on npm View on GIHub View Scorecard		NUMANA FALI NOM FALI	Representations Image: Second Seco	Instance: > doudy-yellow-brownet-washler-4455 Source Anatone: strender Dones: Snaphts: Austriag: Non-45D Logs: Settings Billion Source: Sour	Open Killer (2) A Recent Activity Fel Sage 22 2023 A Non-RED Settings Upstand 115280 Non-RED Settings Upstand
About Grayify This node supports highly-scalable event-driven image conversion to grayscale using the popular ImageMagick software. ImageMagick® is a free and open-source software suite for displaying, converting, and editing		All address states Image: State State State States Image: State State States Image: State State States Image: S	De De DE	Tora Alese Zon	Test Updated Joint From Phylopel Pyre Default / Nodes 802 31.0 Testinger Default / Nodes 802 31.0	10.64 Most PD Service Update PALL PD Service Updat
raster image and vector image files. It can read and write over 200 image file formats, and can support a wide range of image manipulation operations, such as resizing, cropping, and color correction. [ImageMagick](https://imagemagick.org/index.php)		 underweise kann der versten sonnen der	An and an and an and an and an			

Al4 meosc Co-funded by **Al4Compose:** Flowfuse User Journey the European Union



Sing up in the portal

- Wait for admin approval
- OIDC integration desired, currently not possible

Create and/or connect to a Node-RED instance

Inside "Applications"

Use the Node-RED instance:

- Import Flows & Subflows
- **Design your pipelines**
- Execute them

3



Official documentation available in Confluence.

Al4Compose: Innovation in Al4EOSC

- Elyra in <u>EGI Notebooks</u>: we have contributed to have official support of Elyra inside the service, so users can easily access the tool.
- Notebooks to be used in Elyra for AI models of the DEEP Open Catalog:
 - <u>https://github.com/ai4os/ai4-compose/tree/main/elyra/examples</u>
- Documentation in <u>Confluence</u>: soon we will contribute to the official documentation of the project: <u>https://docs.ai4eosc.eu/en/latest/index.html</u>

Raufactorana Q	%
Fitter flas by name	
Notebook	9
P Name A Lat Molfied	
749140 👌 🛃 👌 🕞 👔	· ·
	88
b) distribution distribution di la distributione di la distribu	
🗮 🖿 node_modules 6 months ago	
in services 7 months ago 2. Console	
Concertain and a second s	
🐅 👔 package-lock jaon 6 monthu ago 🟓 👤 🌔 👶 🍘	
Bi (pyranis) uzu z.a. z. U czuwi vytownijstka	
Se Elyra	
WithSteen	
(3350	
	- 11
Handoon File Julia File Show	
Control Info	
	Launcher
	_

de ⊙ Issues រឿ	Pull requests 🕑 Actions 🖽 Projects 🖽 W	Viki ① Security	🗠 Insights 🔞 Setti	ngs		
ai4-compose 🖭	blic)	🖍 Edit	Pins • OUnwatch 5	* § Fork 1	▼ ☆ Star 0	
main - 🐉 1 bran amcaar Merge pull re	ch 🔖 0 tags	Go to file Add	file - <> Code -	About Low-code composi pipelines	ition of AI inference	ŧ
elyra	Merge pull request #10 from Dialdroid/	/dev-dieagra	4 months ago	Readme		
node-red	Update README.md		👌 File Edt View Rum Hermal Edt Tubs Sa	১০০০০০০০০০০০০০০০০০০০০০০০০০০০০০০০০০০০০	se	
LICENSE	Initial commit		The final by cares	٩	2 iarrter × 2 courpipalia Q, 10 > 3 0 2 0 10	Autore Develo E Mont Moranis Motimorates
README.md	Update README.md to add acknowledge	gments		Les Modile s du sp técnique técnique técnique s du sp técnique s du sp		Divisionment Variable (
README.md			(0) E D Anjaray	7 den ap	• A generate_code •	• K occur.compy.
	T		6. 10 3			Britoner Weder Dictor War War War
	Al4 OPec	SC				Environment Vestable* LGER Value

Conclusions



- Al4Compose offers a visual support (drag & drop + customization) to compose Al inference pipelines.
- It minimizes the orchestration effort:
 - Multiple AI models can be triggered for inference and later aggregate the results for enhance accuracy.
- Reusable functions:
 - Pre-defined workflows can be created to facilitate interaction among the AI models in the DEEP Open Catalog.
 - Specific nodes can be created for the different AI models for a simpler definition of workflows.
- Workflows along the computing continuum can be supported (e.g. OSCAR clusters in disparate computing infrastructures).
- Each node can be configured to invoke an OSCAR service within a specific OSCAR clusters.
- Deployment of pre-defined Node-Red instances to facilitate AI composition of workflows (multi-tenant support thanks to FlowFuse).

Bratislava Workshop – 15/11/2023



Thank you! Any questions?

Amanda Calatrava