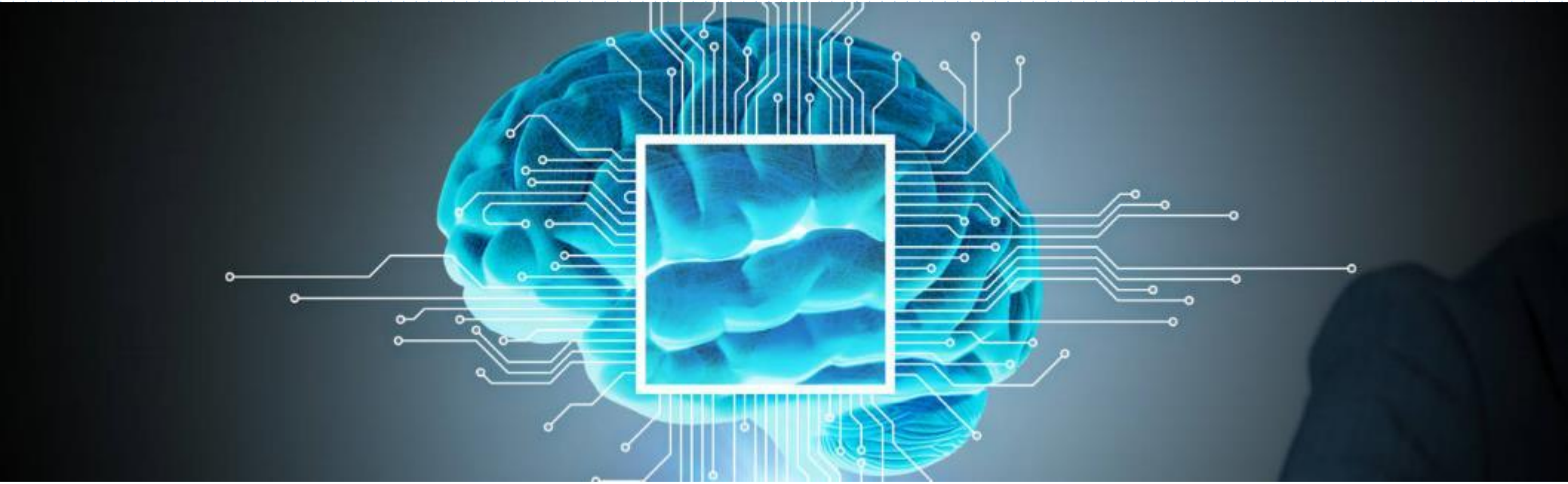
DEEP as a Service



Álvaro López García
<aloga@ifca.unican.es>
IFCA (CSIC-UC)

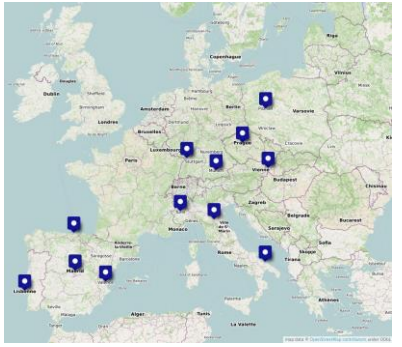
DEEP-Hybrid-DataCloud is funded by the Horizon 2020 Framework Programme of the European union under grant agreement number 777435

DEEP in 1 slide

Designing and **E**nabling **E**-Infrastructures for intensive data **P**rocessing in a **Hybrid ataCloud** (Grant agreement number 777435, Nov 2017 – Apr 2020)

Global objective: Promote the use of **intensive computing services** by different research communities and areas, and the **support by the corresponding e- Infrastructure** providers and open source projects

- Focusing on **Machine learning**, **Deep learning**, and **Post processing** services for the European Open Science Cloud



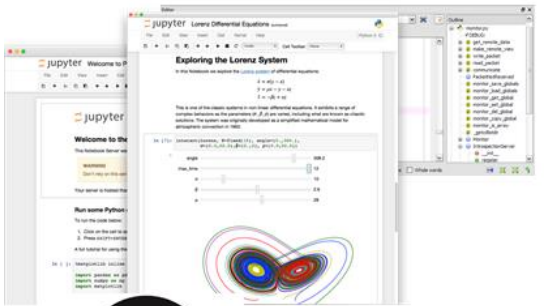
More details: https://indico.egi.eu/event/5000/contributions/14332/

Development cycle in Machine Learning

- The **development of a model starts on a PC**
- The **model is trained** with training data (labelled or not) **and tested** with a different data set
- This phase requires an **adequate computing infrastructure**
- The **model is evaluated** with data that you have never seen before.
- The **cycle is repeated** until the evaluation is satisfactory
- The **work is published** (academy): architecture, data configuration, scientific article, etc.
- Or it is **deployed in production** (industry)

Open issues:

- Access to infrastructure?
- How to share and exchange knowledge?
- Is there a standard way to deploy it as a production service?
- Good practices?
- And many more ...



The goal: Deep as a Service

Training a Machine Learning/Deep Learning model is a very complex and computationally intensive task requiring the user to have a full setup involving a certain **hardware, the adequate drivers, dedicated software and enough memory and storage resources.**

The DEEP-HybridDataCloud project offers a **framework for all users**, and not just for a few experts, enabling the transparent training, sharing and serving of Deep Learning models both locally or on hybrid cloud system.

Covers all the phases of a Machine Learning/Deep Learning project life cycle:

- Development or update of a model→ DEEP Open Catalog
- Model training, test and evaluation→ DEEP training center
- Model deployment as a service → DEEP as a Service
- Model publication for sharing and reusing→ DEEP Open Catalog

DEEP Open Catalog

The **DEEP Open Catalog** provides the universal point of entry to all services offered by DEEP

https://marketplace.deep-hybrid-datacloud.eu

Its offers several options for users of all levels to get acquainted with DEEP

- **Basic Users** can browse the DEEP Open Catalog, download a certain model and apply it to some local or remote data for inference/prediction.
- **Intermediate Users** can also browse the DEEP Open Catalog, download a model and do some training using their own data easily changing the parameters of the training.
- **Advanced Users** can do all of the above. In addition, they will work on more complex tasks, that include larger amounts of data.

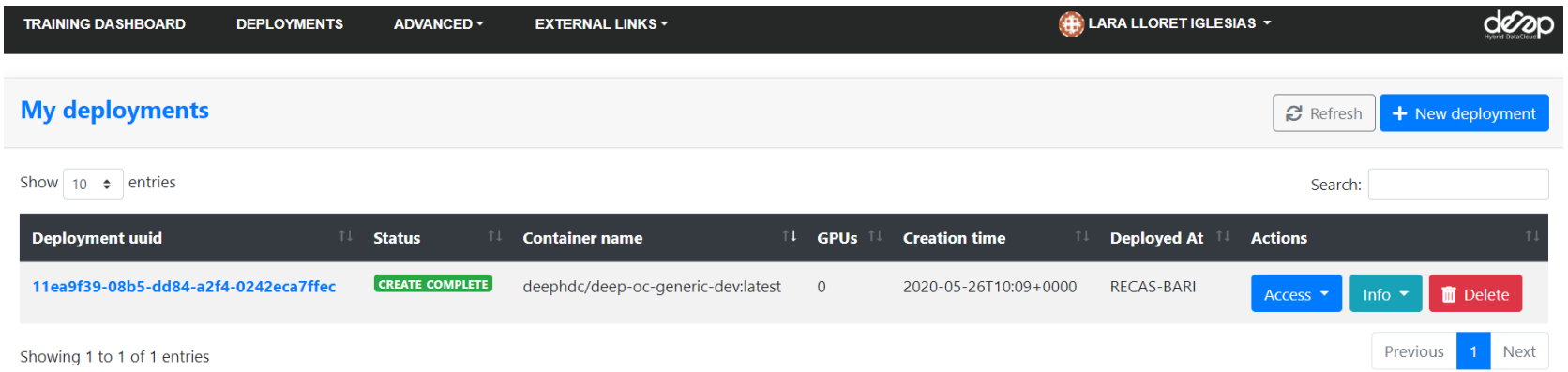
The demo

- When the COVID-19 emergency started we were contacted by the **Marqués de Valdecilla Hospital**
- They were interested in a Deep Learning **image classifier for chest X-Ray images** (pneumonia or non pathological)
- DEEP already had an image classifier in the MarketPlace: **Not adapted for the medical images format (DICOM)**

This demo is based on a real example on how to **adapt an existing DEEP module** for other task, **train** it, **deploy** it as a service and **share** it in the Market place

What you can see in this demo

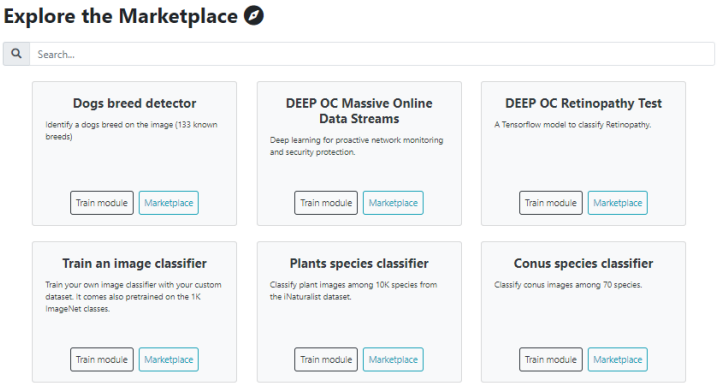
- We will introduce the **DEEP training dashboard** and will deploy the **development module**
  - JupyterLab



- The data (chest x-ray DICOM images) are stored in **OneData**
- We will clone the **adapted repository on image classification** from GitHub and use it for **training with the medical images** → Transfer Learning

What you can see in this demo

- When the new module is trained we want to **include it in the DEEP Marketplace**
- As soon as we commit the code to GitHub the **Jenkins pipeline is triggered** starting the different **software quality tests**
- When the **Jenkins pipeline is finished**, **the module will be available as service the DEEP Marketplace and can be used as a service to predict on new images**



In summary

- **Demo video is available here: https://tinyurl.com/yyuznjdw**
- **We have taken a module from the DEEP market place and adapted it to perform a new task**
- **We have used this new module to train on a new dataset and we have integrated the trained model into the DEEP Market Place**
- **Now, the module is available as a service so that anybody can use our training model to perform some prediction**

Thank you!

Hybrid DataCloud

https://deep-hybrid-datacloud.eu

DEEP-Hybrid-DataCloud is funded by the Horizon 2020 Framework Programme of the European union under grant agreement number 777435