



DEEPHEALTH

DeepHealth perspective

HPC, Big Data, IoT and AI future industry-driven
collaborative strategic topics

Workshop at the EuroHPC Summit Week 2021

Monica Caballero (Project Manager)

monica.caballero.galeote@everis.com

Jon Ander Gómez (Technical Manager) – jon@upv.es

March 23rd, 2021

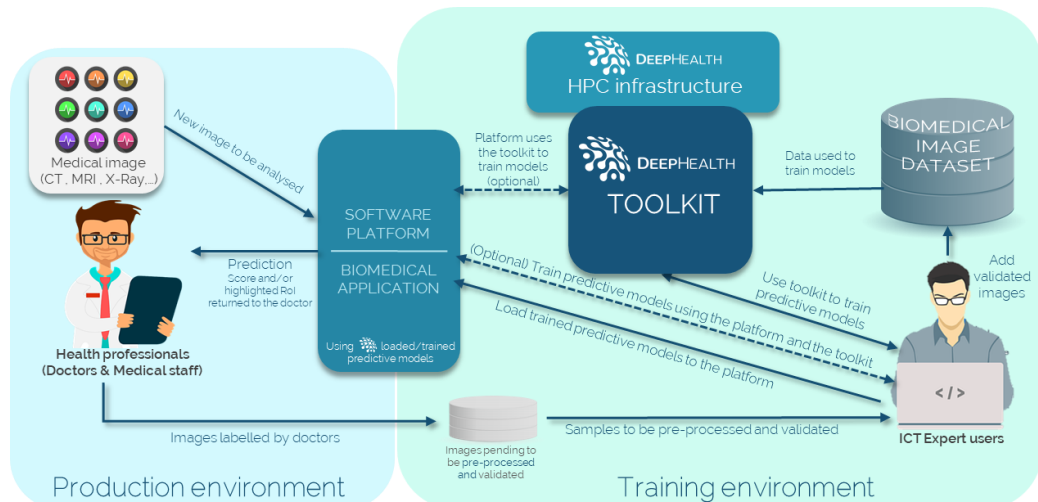


The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825111.



About DeepHealth

Main Goal : Put HPC computing power at the service of biomedical applications with DL needs and apply DL techniques on large and complex image biomedical datasets to support **new and more efficient ways of diagnosis, monitoring and treatment of diseases.**



- **The DeepHealth toolkit**
 - EDDL: The European Distributed Deep Learning Library
 - ECVL: the European Computer Vision Library
- **HPC infrastructure for an efficient execution of the libraries**
- **Seven enhanced biomedical and AI software platforms**
- **Validation in 14 healthcare use cases**

Research Organisations



Health Organisations



Large Industries

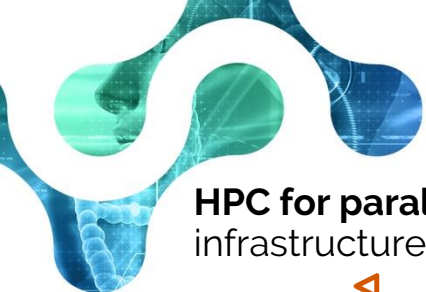


SMEs



Key facts

	Duration: 36 months		Budget: 14.642.366 €		22 partners from 9 countries:
	Starting date: Jan 2019		EU funding: 12.774.824 €		Research centers, Health organizations, large industries and SMEs



HPC in use cases, Challenges and Plans

Data / Workflows / HPC-Cloud infrastructure / AI-ML training



DEEPHEALTH

HPC for parallelizing the training operations of AI/ML use-case models on top of HPC infrastructures for reducing training times and enabling inference on embedded systems.



Impact- Challenges



DeepHealth contributions

Data

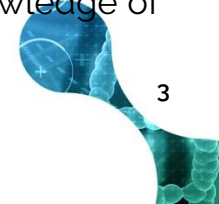
- Getting **health data out of health institutions**-premises
- Quantity, quality and Shareable-interoperable data for AI purposes.

- Definition of a **data-lake structure** and organization, and anonymization procedures for health data.
- Exploration of **federated learning techniques** to avoid the need of moving/centralize the data

Workflows

- No established workflows combining DL+BigData+HPC+Cloud+CV+Health
- Lack of highly-skilled experts in all these areas to leverage HPC for DL and CV purposes.
- Need for portable solutions that industry can easily adopt.

- **Portability:** AI+HPC workflows for training & inference relying on task-based programming models and hybrid cloud-HPC cross-application workflows to efficiently exploit HPC-Cloud infrastructures
- **Usability:** DeepHealth toolkit to ease the work of computer/data scientist with no deep knowledge of DL and HPC management





HPC in use cases, Challenges and Plans

Data / Workflows / HPC-Cloud infrastructure / AI-ML training



DEEPHEALTH



Impact- Challenges



DeepHealth contributions

HPC-Cloud infrastructure

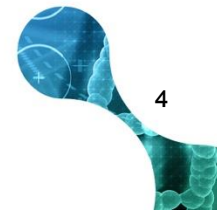
- Transparent use of heterogenous HPC (featuring CPU, GPUs, FPGAs, and other HW accelerators) and Cloud infrastructures
- Support flexible computational requirements

- Provide an HPC infrastructure for an efficient execution of the libraries
- **Transparent:** Abstract the parallel execution from the underlying infrastructure.
- Promote a **"cloudified approach"** to HPC

AI-ML training

- Training time when dealing with images
- Great variety of available tools, which one to use? No EU resources
- DL and CV to work together efficiently.

- DeepHealth toolkit including ECV and EDDL libraries ready to run on HPC/Cloud infrastructures in a transparent way for computer/data scientist working in the Health sector, or any other sector





Prioritization of the four fields

In terms of complexity and importance for R&I calls in Europe



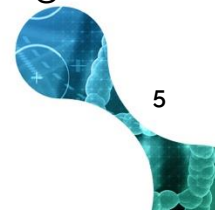
DEEPHEALTH

1) Data

2+3) HPC + Workflows

4) AI/ML

- The availability of FAIR **data** is still a big challenge:
 - Difficult to make data providers from the same sector (e.g. Health sector) to collect data following **standard protocols** (still to be defined in most of cases) to make datasets corresponding to the same disease collected from different hospitals to be **interoperable** to use them together train AI/ML models.
 - Boost data-sharing in some domains, specially in the health sector (limiting the exploitation of available data)
- **HPC & (AI+HPC) workflows** needed to be boosted to increase the productivity of expert-users (data-scientists)
- **AI/ML**: it is a mature enough research area. But still a long way to go regarding the improvement of model accuracy. Also explainable AI.





Industry in shaping future HPC strategy

Unique HPC needs of industrial partners (IT partners serving the Health industry)

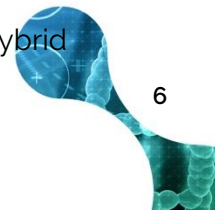
- **Time-to-solution: reducing processing times** for **incorporating AI/ML predictive models to their applications and platforms** to solve health use-cases to support the diagnosis, treatment and monitoring of diseases
- **Easy to use:** if properly engineered (e.g. cloudified), HPC is highly desired to allow easy update of AI/ML models to adapt to new use-cases and improve models fast with new available data

How do you think that industry is engaged to the above-mentioned areas?

- **Expectancy on how they can benefit from HPC** technologies in their AI strategy, applications and services and **demanding** data, workflows, AI/ML tools and pre-trained models.
- Most industrial partners have only temporary needs of high processing power (generate the model, update it), thus HPC solutions provided **as a service (e.g. cloudified HPC)**, or low-power (e.g. FPGA-based) inference for embedded systems could be of interest for them

What are your ideas about a commercialization of the product results?

- The **DeepHealth toolkit** is conceived as free and open-source software available on a public repository, with a sustainability plan based on *services and advice to any company or academic institution interested in using any of the software components.*
- **HPC+cloud results**, commercial exploitation for different results by industrial partners developing FPGA and hybrid cloud solutions.





DEEPHEALTH Questions?

Project Coordinator: Mónica Caballero monica.caballero.galeote@everis.com

Technical Manager: Jon Ander Gómez jon@upv.es

<https://deephealth-project.eu>



@DeepHealthEU

