GENCI

Le calcul au service de la connaissance

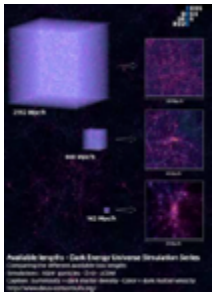
Overview of converged applications

Stéphane REQUENA, GENCI

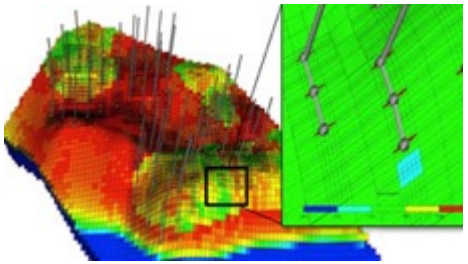
TECHNOLOGICAL CONTEXT

Convergence HPC / DATA / AI

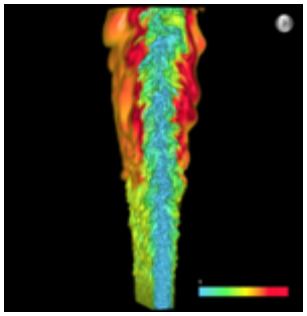
❑Explosion of computational data



Cosmology
DEUS project
150 PB raw data



Reservoir modelling
of giga models 350 TB/run



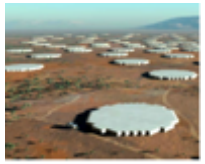
HiFi turbulent
DNS combustion
S3D : 1PB / 30mn

Climate CMIP exercises

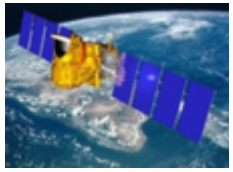
Status CMIP5 data archive:
- **1.8 PB for 59000 data sets stored in 4.3 Mio Files in 23 ESGF data nodes**
- **CMIP5 data is about 50 times CMIP3**

Extrapolation to CMIP6:
- **CMIP6 has a more complex experiment structure than CMIP5.**
- **Expectations: more models, finer spatial resolution and larger ensembles**
- **Factor of 20: 36 PB in 86 Mio Files**
- **Factor of 50: 90 PB in 215 Mio Files**

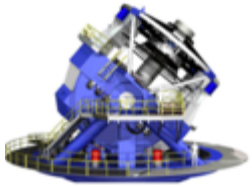
❑And instrumental data as well



LOFAR/SKA
16 TB/s raw

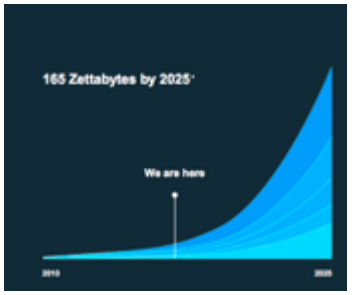


COPERNICUS/SWOT
4 PB/day raw



LSST/EUCLID
20 TB/night raw

Seismic sensors
100 TB/yr



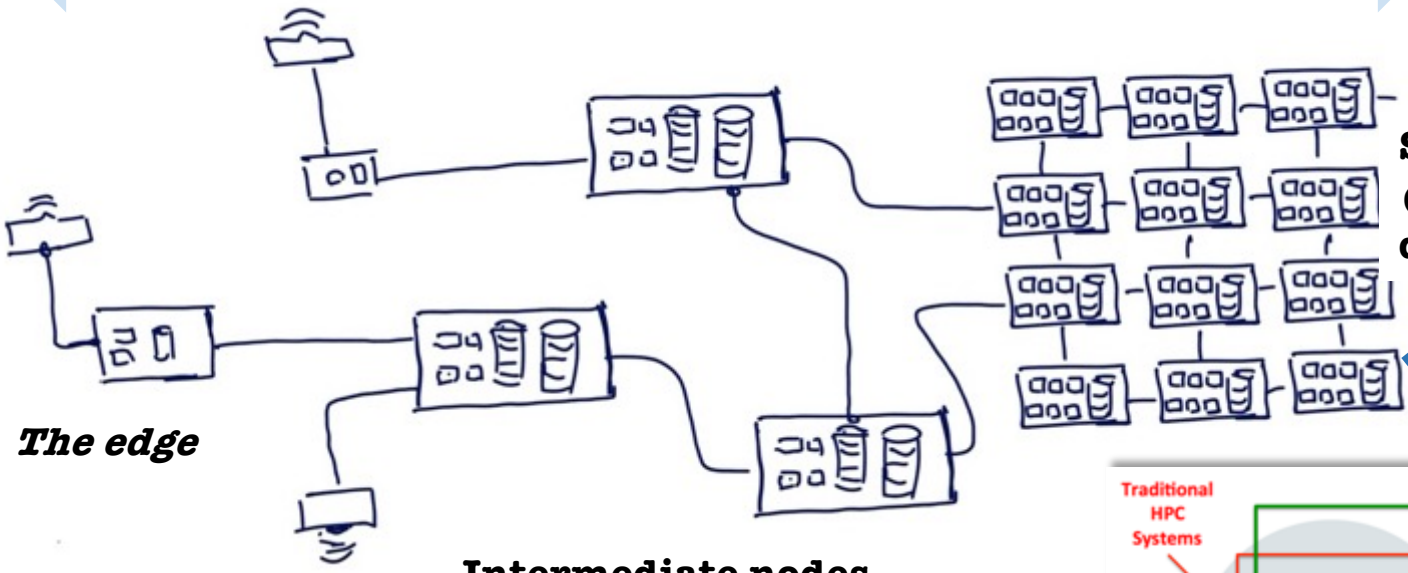
Internet & **IoT**

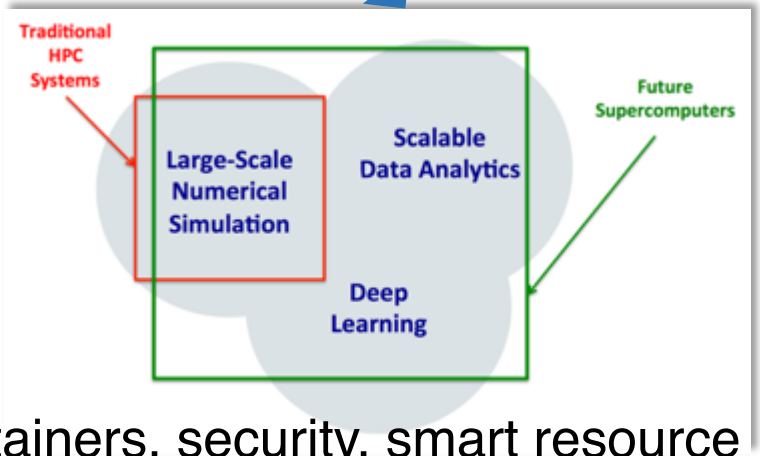
First picture of black
hole (M87) = 15 PB

TECHNOLOGICAL CONTEXT

Toward Advanced Cyberinfrastructure Platforms

a new digital continuum

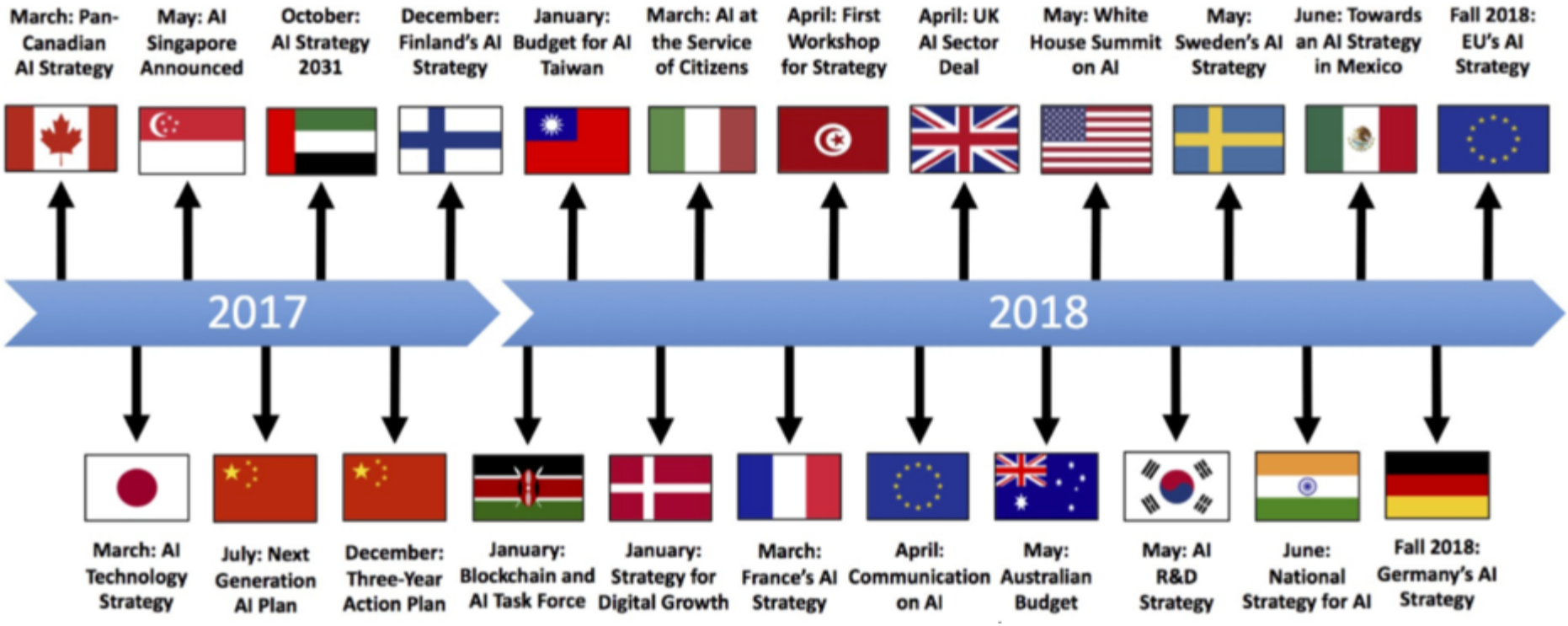




- Challenges : cohabitation of SW stacks, containers, security, smart resource managers, end to end workflows (from the edge to the tape), …
- Development of **new services, co design** and **user support**

CONVERGENCE HPC – AI

AI roadmaps since 2017



HPC/AI CONVERGENCE HAS STARTED

Already since few years in Japan, USA and EU

Tsubame3

ABCI

Fugaku

Summit

Aurora

Frontier

El Capitan

Marconi100, CINECA

Jean Zay, GENCI

Juwels, FZJ

PizDaint, CSCS

LUMI

EuroHPC

Leonardo + Future exascale systems

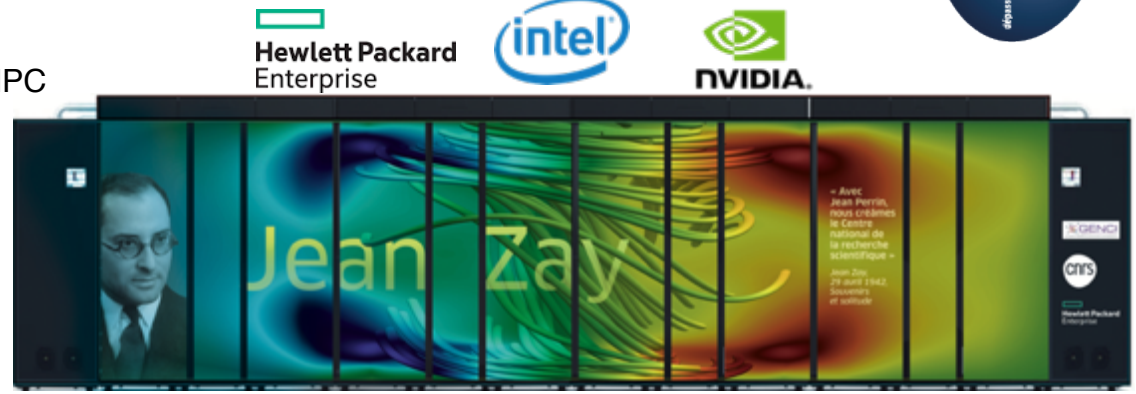
MN5

INTRODUCING JEAN ZAY @ IDRIS

One of the biggest converged systems in Europe

❑ Objectives

- **Support** with **sovereign** and **leading edge** HPC facilities the French AI research community
- **Foster** synergies between AI and HPC communities
- To be **integrated** into the French AI plan



❑ **Converged** system

- HPC + HPDA + AI

❑ New dynamic access modes

- Fast access to elastic pool of resources
- Support of containers, notebooks, …

❑Big Milestones

- March 2018 : French "**AI for Humanity**"
- Jan 2019 : Contract between HPE and GENCI announced
- May 2019 : Installation
- October : 2019 Full production
- *Mid 2020 : Major upgrade (*2)*

28 PF

❑A balanced compute

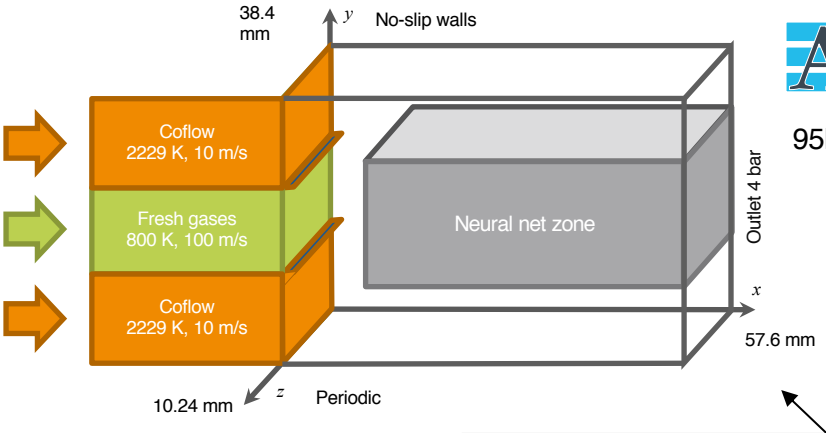
- HPE SGI 8600
- Scalar partition (HPC): 1528 nodes, 3056 CPU CSL 6248, **61 120 cores**, OPA
- Converged partitions – **2696 GPU**
  - 612 thin nodes, **2448 GPU V100 16/32GB**, 4xOPA
  - 31 fat nodes, **248 GPU V100 32 GB**, 4xOPA

❑ and storage architecture

- **2.2 PB @ 0.5 TB/s** full flash (L1)
- **35 PB @ 150GB/s** HDD (L2)
- SpectrumScale parallel filesystem

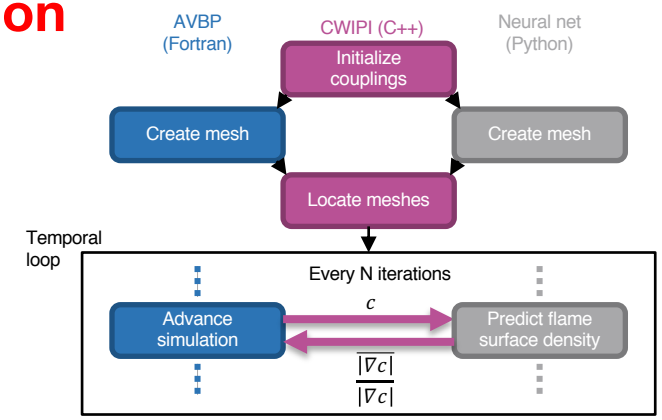
SOME EARLY RESULTS

HPC and AI for high fidelity combustion

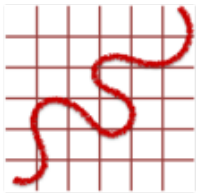


AVBP

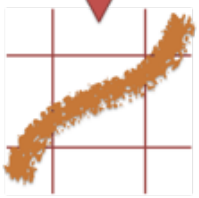
95M cells



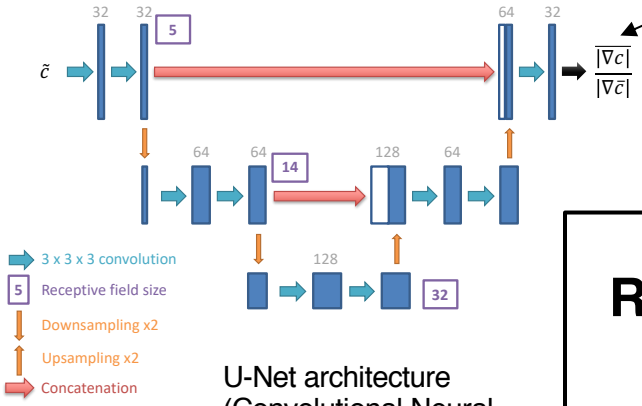
Large-scale LES of a slot jet flame **coupled** with a neural network-based wrinkling model



DNS: resolved flame



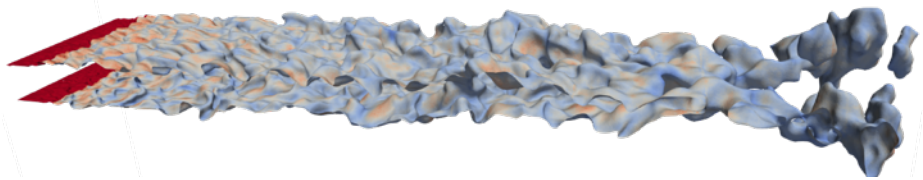
LES: filtered flame



U-Net architecture (Convolutional Neural Network)

Computing the total surface of the flame front is very important in combustion simulations. In LES this requires a model for unresolved subgrid-scale wrinkling

Results



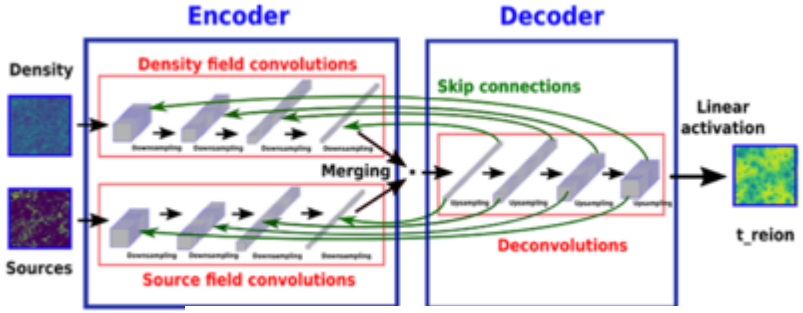
- Successful co-simulations **up to 2816 CPUs and 256 GPUs**
- Dynamic behaviour of the flame is well-recovered
- Major improvements in the approach were made possible by the analysis of the model's behaviour on this complex configuration

SOME EARLY RESULTS

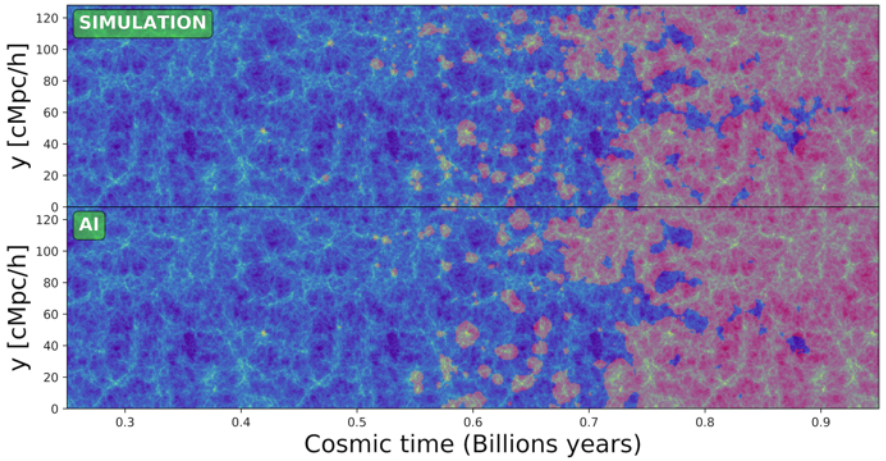
Coupling learn models and simulation models in cosmology

Observatoire astronomique de Strasbourg

- Principal Investigator : D. Aubert from Obervatoire de Strasbourg
- Study of the **reonization** of the Universe = 1 Gyrs after Big Bang
  - Will be observed soon by instruments like EELT, JWST or SKA
  - For the moment only based on massive and costly simulations coupling gravitation, hydrodynamics and radiative transport
  - Idea : couple gravitation/hydrodynamics numerical models AND learnt radiative transport models



  - Use of auto-encoders based on TensorFlow and Keras
  - Methodology already validated for small cubes of 128 Mpc/256^3
  - Target = 128 Mpc on meshes of 1024^3



SOME EARLY RESULTS

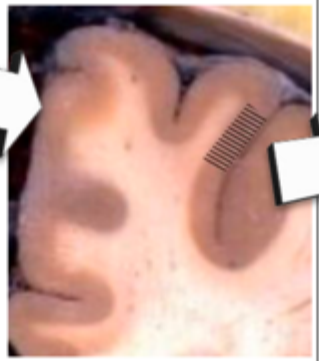
Using simulation and AI for understanding brain functions

Learning the diffusion MRI signature of brain cytoarchitecture using AI & large scale simulations to decipher the anatomy of brain functions - C. Poupon & K. Ginsburger

human brain

Functional Networks (using fMRI)

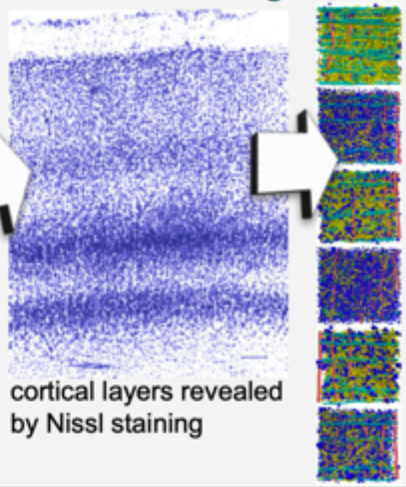
Activation of visual functional areas in the occipital lobe



cerebral cortex populated by neurons and glial cells

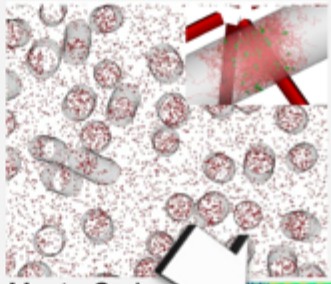
link

Large scale simulation of realistic cortical layers

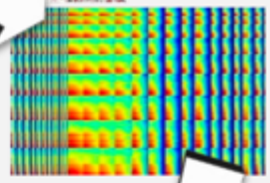


cortical layers revealed by Nissl staining

Large scale simulations of virtual cortical layers

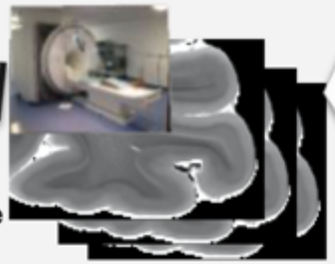


Monte-Carlo simulations of Diffusion MRI Signatures for each virtual cortex sample



In vivo cytoarchitectonics mapping

In vivo mapping of the human cortex cytoarchitecture using the decoder from MRI acquisitions



Training a Machine Learning decoder

Use of Machine Learning / AI tools to learn the link between the Diffusion MRI signature and the cortex cytoarchitectural features

SOME EARLY RESULTS

ADAGE: computer-Aided Diagnosis of AGE-related brain diseases

Principal investigator: Olivier Colliot, ARAMIS Lab (CNRS, Inria, Inserm, Sorbonne, ICM)

Develop and validate deep learning tools for diagnosis from very large scale medical imaging data

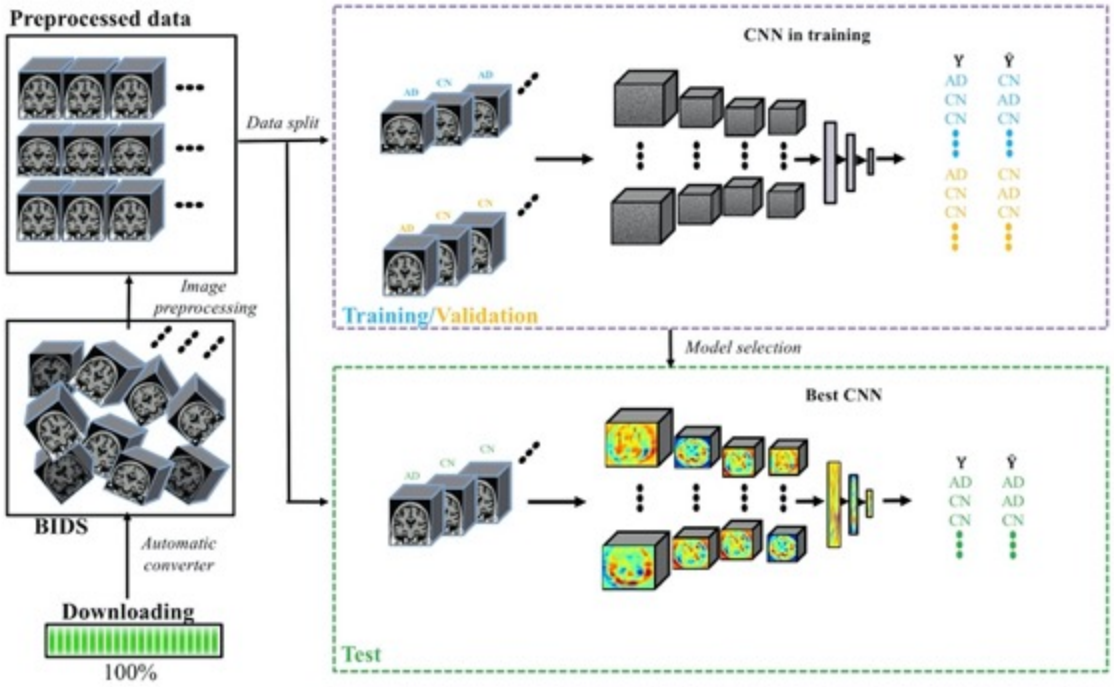
Over 20,000 patients
Up to 128 GPUs used

Results:

- prediction of Alzheimer's with high accuracy

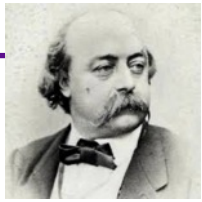
Ongoing:

- model of ageing
- diagnosis of other pathologies



COCORICO HERE IS FLAUBERT

A(nother) French NLP model



- Developed using PyTorch by researchers from LIG, LAMSADE et LLF
- French version derived from initial BERT (transfer training)
- Final training using a 71 GB corpus of texts with 137M parameters
  - Wikipedia FR, journal Le Monde (1987-2003), books from Gutemberg project
  - (French) Minutes of debates of European parlement
- Use of Jean Zay (IDRIS/GENCI) on >800 GPU (128 initialy)
  - «I*t was necessary to jump on this unique opportunity to create a French language resource of this scope. To do this, we set up the FlauBERT team to exploit this essential computing power of Jean Zay but which was inaccessible to us and unfortunately reserved until now for GAFAM.* » A. Allauzen (LAMSADE)
- Concurrent project called CamemBERT (Facebook, Inria) but trained with half the data

https://github.com/getalp/Flaubert

Model	Books	DVD	Music
MultiFiT†	91.25	89.55	93.40
mBERT†	86.15	86.90	86.65
CamemBERT	**93.40**	**92.70**	94.15
FlauBERT$_{BASE}$	**93.40**	92.50	**94.30**

† Results reported in (Eisenschlos et al., 2019).

Table 3: Accuracy on the CLS dataset for French.

GENCI

Le calcul au service de la connaissance

Thank you!

Stéphane REQUENA, GENCI