LightOn
We bring Light to AI

Unlocking Transformative AI with Photonic Computing

Laurent DAUDET

LightOn, CTO and Co-Founder

SwissPhotonics Workshop on Optical Computing –April 14th, 2021

4 Co-Founders in a team of 25

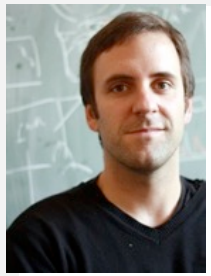
Igor Carron

CEO

Laurent Daudet

CTO

Professor at
Université de Paris
(on leave)



Sylvain Gigan

Professor at
Sorbonne University
OPTICS advisor

Florent Krzakala

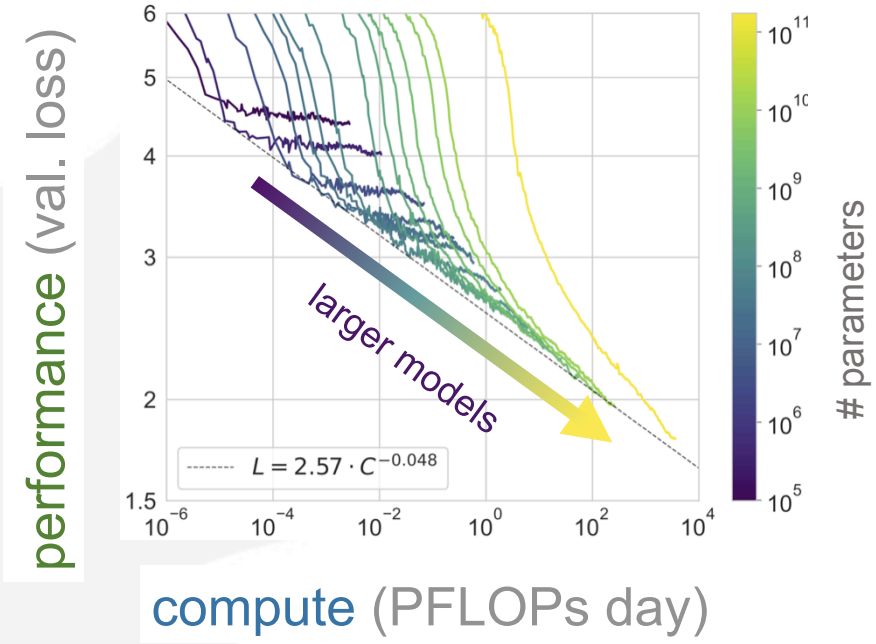
Professor at EPFL
MACHINE LEARNING
advisor

Founded 2016
Based in Paris, France

The AI scaling hypothesis

Scaling laws in Language Models [Kaplan *et al.*, 2020]

With strong and scalable priors, increased model size is all you need !



Larger models score higher, generalize better, train faster

> The biggest lesson that can be read from 70 years of AI research is that general methods that leverage computation are ultimately the most effective, and by a large margin.

Rich Sutton, The Bitter Lesson.

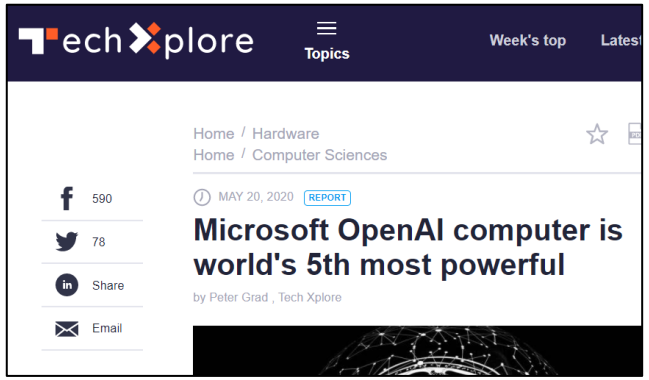
Reaching the limits ?

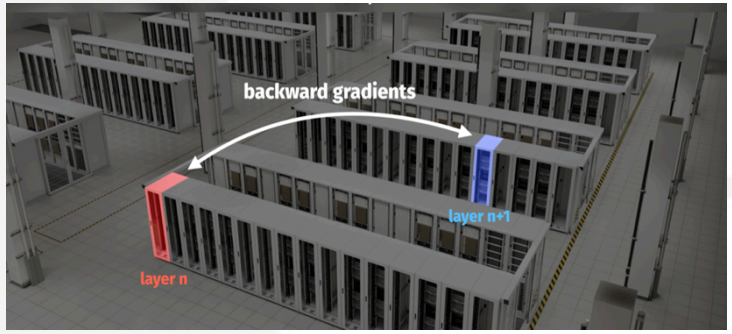
Impressive performance boost of GPT-3 over GPT-2 in only a year !

Same Transformer-based model

....but **BIGGER**

- Amount of training data x50 – « just » linear
- Number of parameters **x110**





The challenge: **Superlinear Scaling** !

++ raw compute

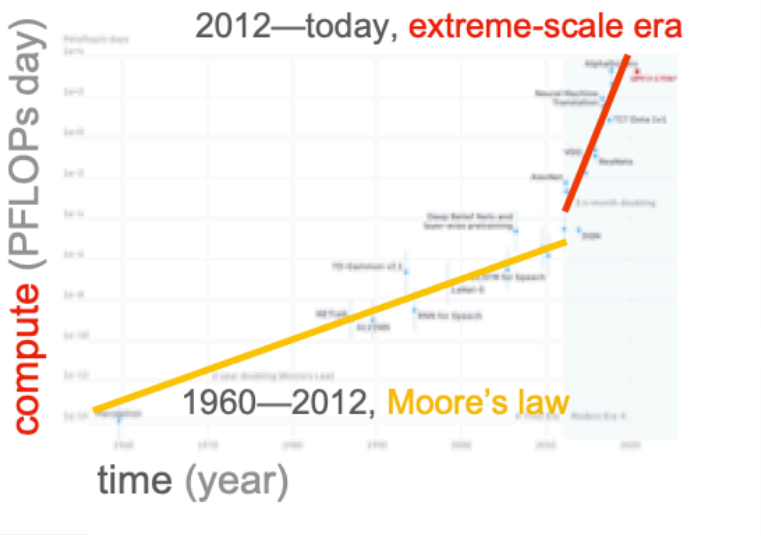
++ memory

++++ communication between compute nodes

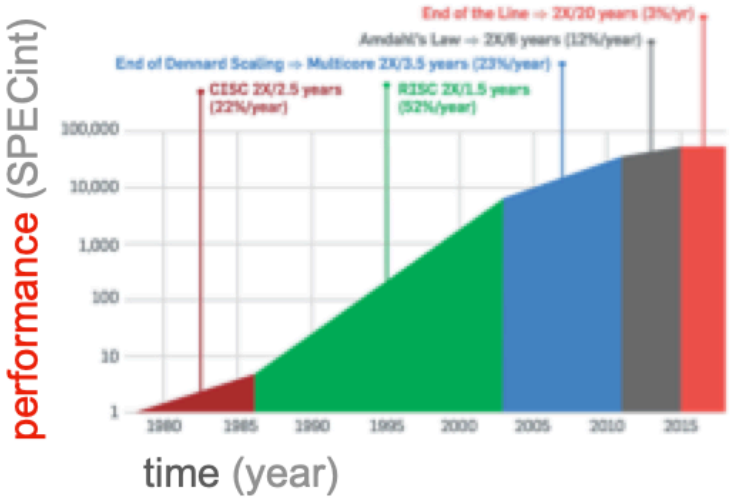
Scaling laws → training **a truly scaled-up « GPT-4 »** (10 trillion parameters) would require a 8,000 A100 GPUs supercomputer **running for 17 years**

Transformative AI will not happen with silicon only

How to reconcile deep learning gargantuan needs with diminishing hardware gains ?



Modern deep learning significantly outpaces Moore's Law



Hardware progress is increasingly bound to specialization

The computing hardware problem is an algorithm problem

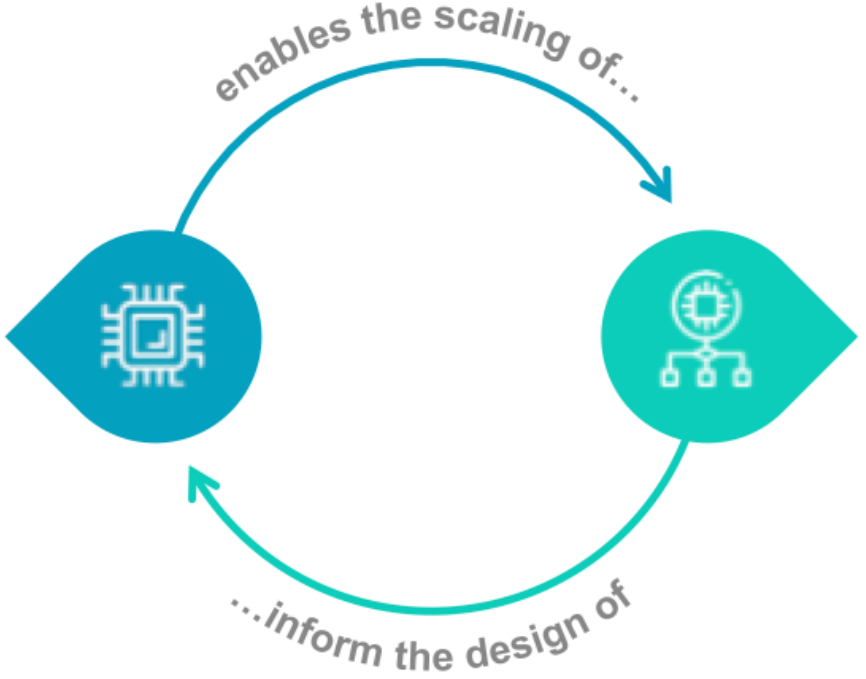
The Hardware Lottery

Sara Hooker — August, 2020

“How does tooling choose which research ideas succeed and fail, and what does the future hold?”

Hardware *at the speed of software*

PHOTONIC HARDWARE



IMPROVED ALGORITHMS

What is LightOn OPU ?

LightOn OPU is the world's first million-scale photonic co-processor dedicated to massive scale AI models

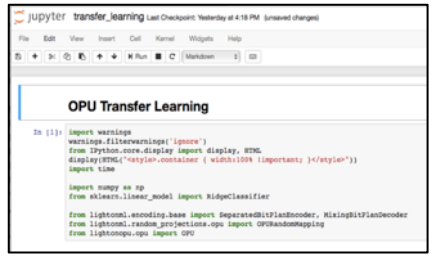
CPU,GPU,
TPU, IPU, etc…

+

Disruptive electro-photonics hardware
OPU: Optical Processing Unit with Photonic Cores

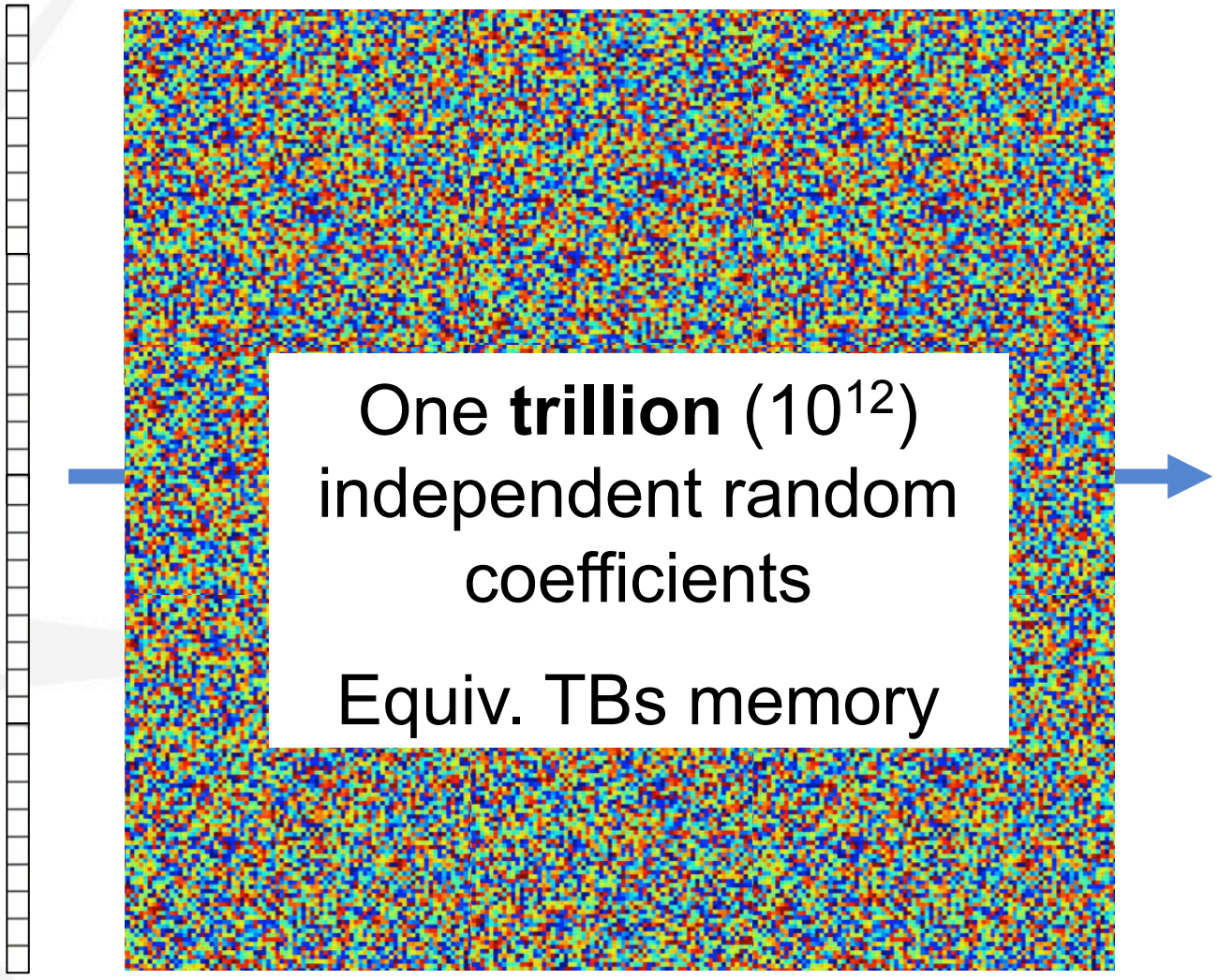
+

Optimized software platform
compliant with Python and ML packages



Matrix-vector multiplication through light scattering

1 million independent input pixels



1 million independent output pixels

Designed for Large Scale Machine Learning

The OPU performs **Random Projections** in the analog domain

input vector x → *output vector* y = Hx

with H a complex random iid matrix

Analog - non Von Neumann − operations

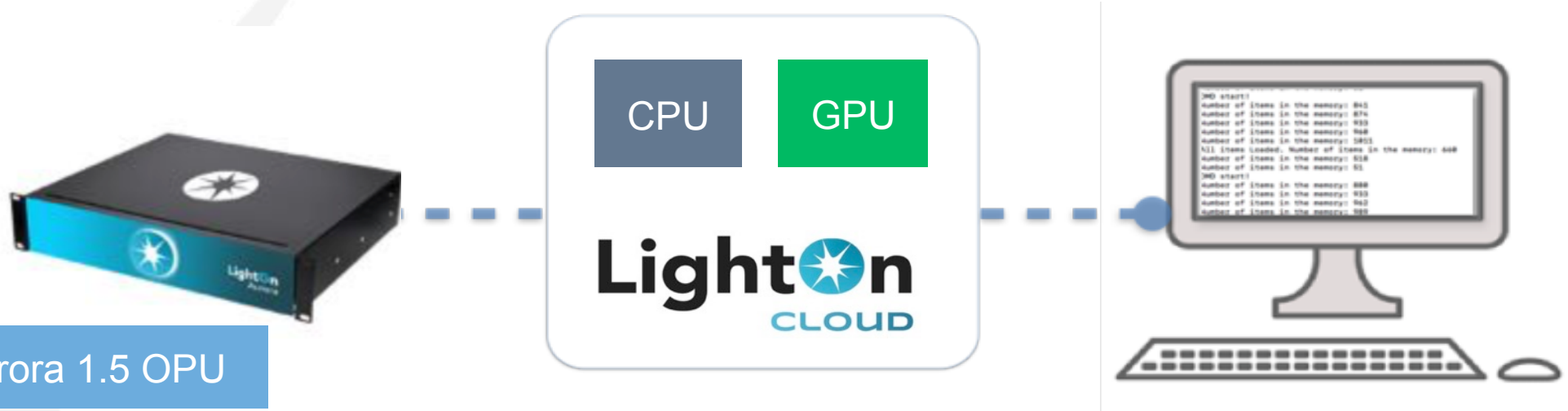
Introducing LightOn Appliance

World's first commercially available photonic co-processor for AI and HPC

on-premises rackable 2U unit, external PCIe connected
1,500 TeraOPS at 30 W TDP

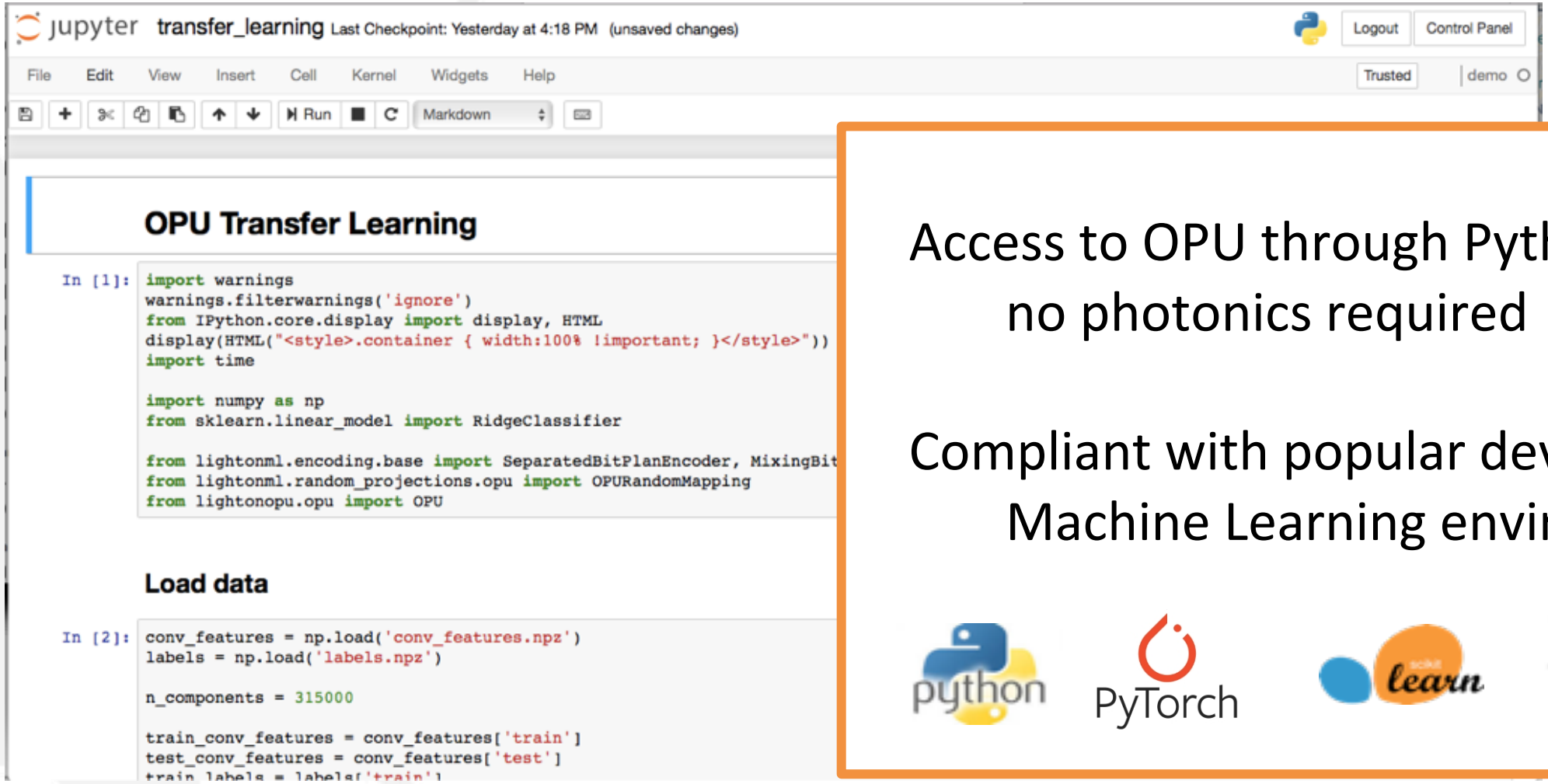
Spec sheet and leasing packages at https://lighton.ai/lighton-appliance/

Test drive *today* on the LightOn Cloud



- Hybrid cloud platform hosted in 2 datacenters in Paris area – contiuously upgraded since Dec 2017
- Researchers get free compute time through **LightOn Cloud for Research** program
- See all info on https://cloud.lighton.ai

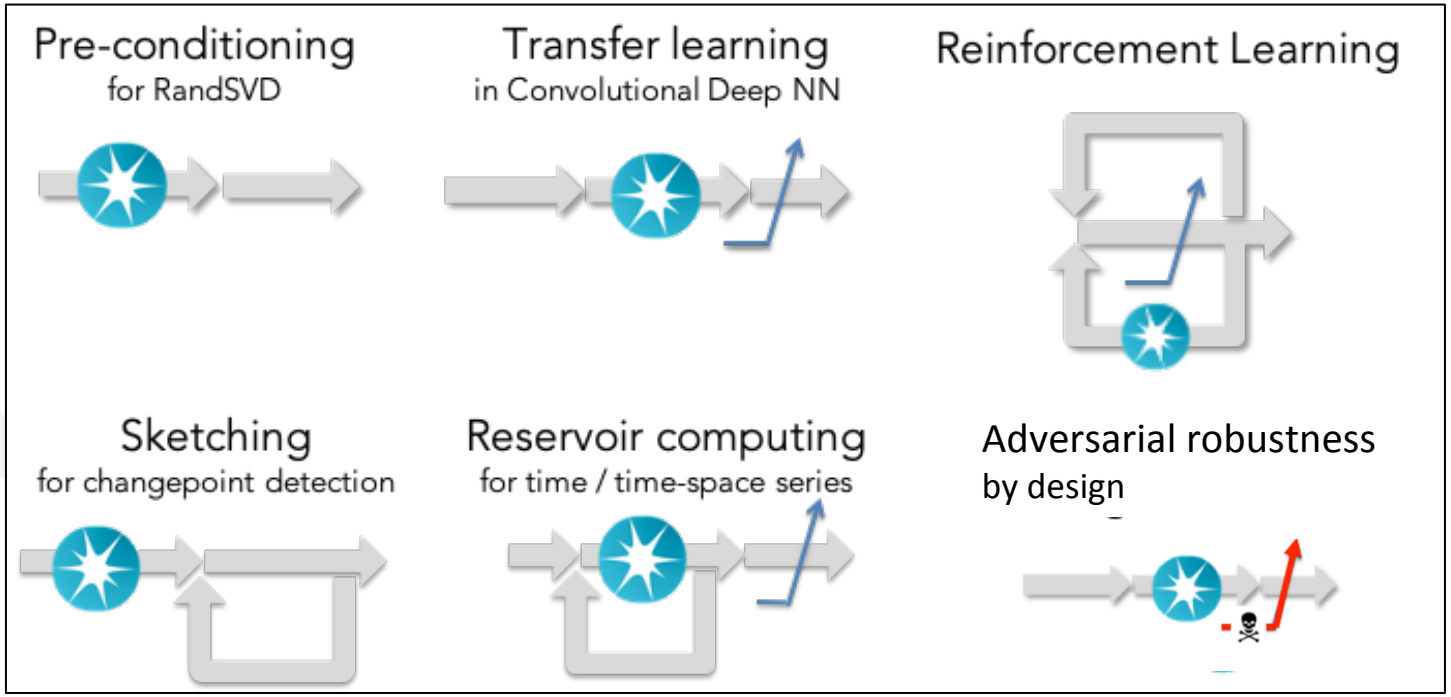
User interface (Jupyter)



Access to OPU through Python library:
no photonics required !

Compliant with popular dev and
Machine Learning environments

Hybrid computing in AI pipelines



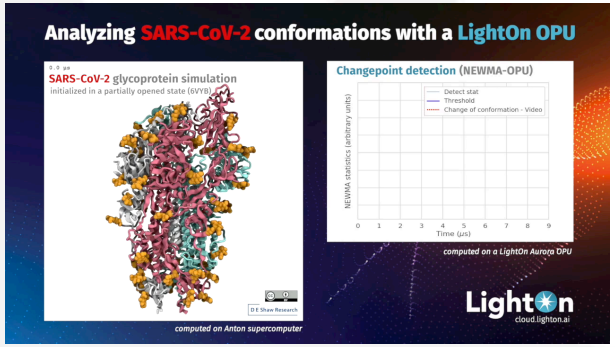
Hybrid computing in AI pipelines

Ivan Dokmanic
Associate Prof.
UNIVERSITY OF ILLINOIS

NEURAL INFORMATION PROCESSING SYSTEMS
NeurIPS 2019

David Rousseau @dhpmrou · 4 avr.
Our talk on analysing #hep data with random matrices in @LightOnIO Optical Processor Unit accepted at @ichep2020 conference ! (Remote in July or Prague beg 2021)#hepml @Laurent_Daudet @IgorCarron

6 15



Pre-conditioning
for RandSVD

Transfer learning
in Convolutional Deep NN

Reinforcement Learning

Sketching
for changepoint detection

Reservoir computing
for time / time-space series

Adversarial robustness
by design

Accelerating SARS-COv2 Molecular Dynamics Studies with Optical Random Features

Amélie Chatelain
LightOn ML R&D engineer

Collaboration with FAIR
Facebook AI Research

NEURAL INFORMATION PROCESSING SYSTEMS
NeurIPS 2020

A new paradigm for AI training

BEYOND BACKPROPAGATION: A NEW DISTRIBUTED TRAINING PARADIGM

NEURAL INFORMATION PROCESSING SYSTEMS

NeurIPS 2020

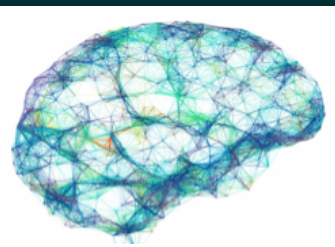
Direct Feedback Alignment Scales to Modern Deep Learning Tasks and Architectures

Julien Launay[1,2] Iacopo Poli[1] François Boniface[1] Florent Krzakala[1,2,3]
[1]LightOn [2]LPENS, École Normale Supérieure [3] IdePhics, EPFL

- **Architecture agnostic**: scales to modern deep learning architectures
  neural view synthesis, NLP, recommender systems ...
- First **optical training** demonstrated on **graph neural networks**
  Oral presentation at NeurIPS 2020 “Beyond backprop” workshop
- Inference on silicon → model portability

At NeurIPS 2020, researchers proposed faster, more efficient alternatives to backpropagation

Kyle Wiggers
@Kyle_L_Wiggers
December 16, 2020 11:18 AM



VentureBeat, Dec 2020

Scaling up

Scaling up jointly in different dimensions:

- The size of Machine Learning models - above 1M parameters models
- The Hardware potential to go x100 in #OPS
  towards ExaOPS Non von Neumann at less than 1 kW
- The use cases: opening to RandNLA (RandSVD → recommender systems, inverse problems...)
- Combining different building blocks with Neural Architecture Search

Building a community

Technical articles, Blogposts, GitHub,
White paper, Newsletter, Meetups ...

Request for access to LightOn Cloud

LightOn.ai

contact@lighton.io

@LightOnIO