Deep Learning Acceleration of Progress in Fusion Energy Research

William Tang
Princeton University/Princeton Plasma Physics Laboratory (PPPL)

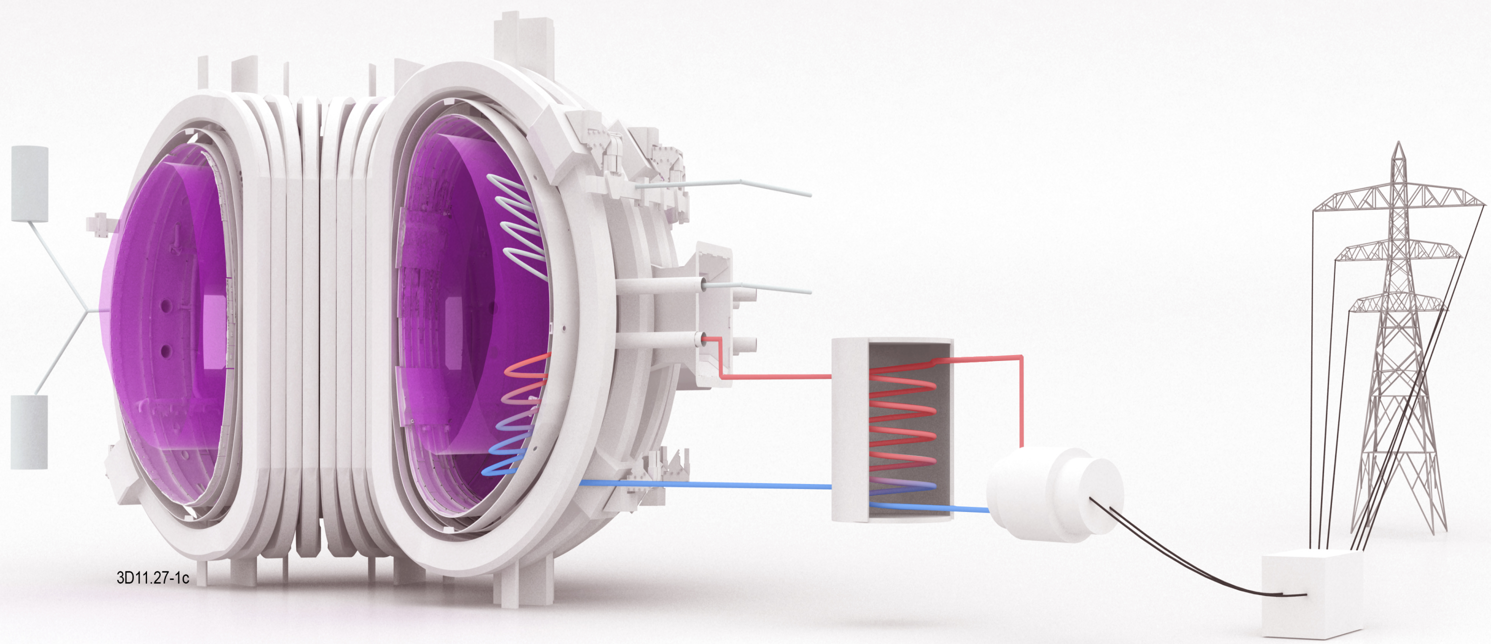
Fast Machine Learning for Science Workshop
Southern Methodist University, Dallas, TX

December 2, 2020

FRNN Project Team

Julian Kates-Harbeck (Kernel Inc./PU), Alexey Svyatkovskiy (Microsoft/PU),
Ge Dong (PPPL, PU), Kyle Felker (ANL/PU), Keith Erickson (PPPL), Dan Boyer (PPPL)

MISSION: Delivery of Magnetic Fusion Power



CNN's "MOONSHOTS for 21st CENTURY" (Hosted by Fareed Zakaria)

– Five segments (Spring, 2015) exploring "exciting futuristic endeavors in science & technology in 21st Century"

(1) Human Mission to Mars

(2) 3D Printing of a Human Heart

(3) Creating a Star on Earth: Quest for Fusion Energy

(4) Hypersonic Aviation

(5) Mapping the Human Brain

"Creating a Star on Earth" → "takes a fascinating look at how *harnessing the energy of nuclear fusion reactions may create a virtually limitless energy source.*"

Stephen Hawking: (BBC Interview, 18 Nov. 2016)

" I would like nuclear fusion to become a practical power source. It would provide an inexhaustible supply of energy, without pollution or global warming."

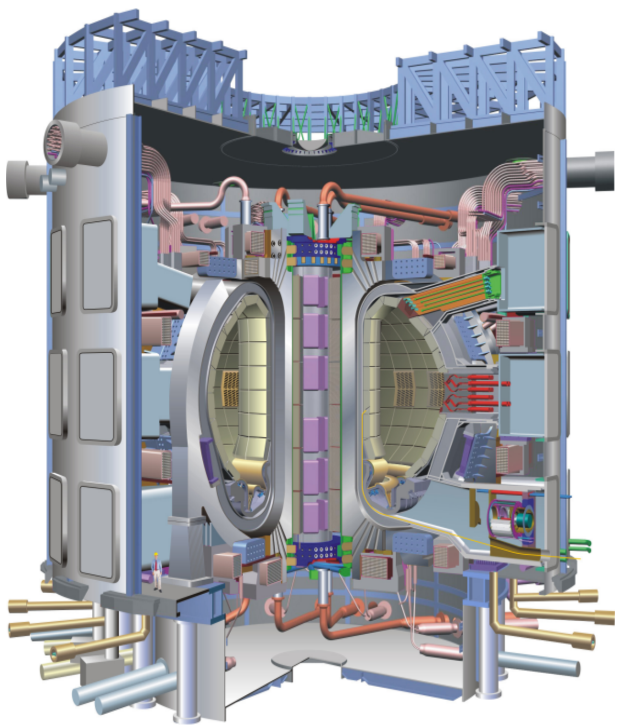
ITER Goal: *Demonstration of the Scientific and Technological Feasibility of Fusion Power*

- ITER *is an ~$25B facility located in France & involving 7 governments representing over half of world's population*
  → *dramatic next-step for Magnetic Fusion Energy (MFE) producing a sustained burning plasma with "1st plasma" targeted for ~ Dec. 2025*

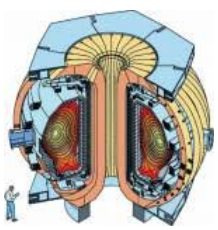
  -- JET: 10 MW(th) for 1 second with gain ~1 (EU)
  -- ITER Goal: 500 MW(th) for >400 seconds with gain >10

- Ongoing R&D programs worldwide [experiments, theory, computation, and technology] *highest priority to predict & control disruptions for ITER*

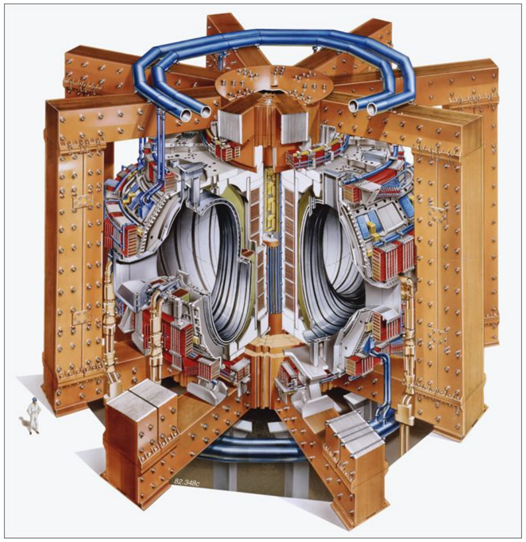
➔ *Realistic AI/HPC-enabled simulations required to cost-effectively plan, "steer," & harvest key information from expensive (~$1M/long-pulse) ITER shots to enable DEMO*



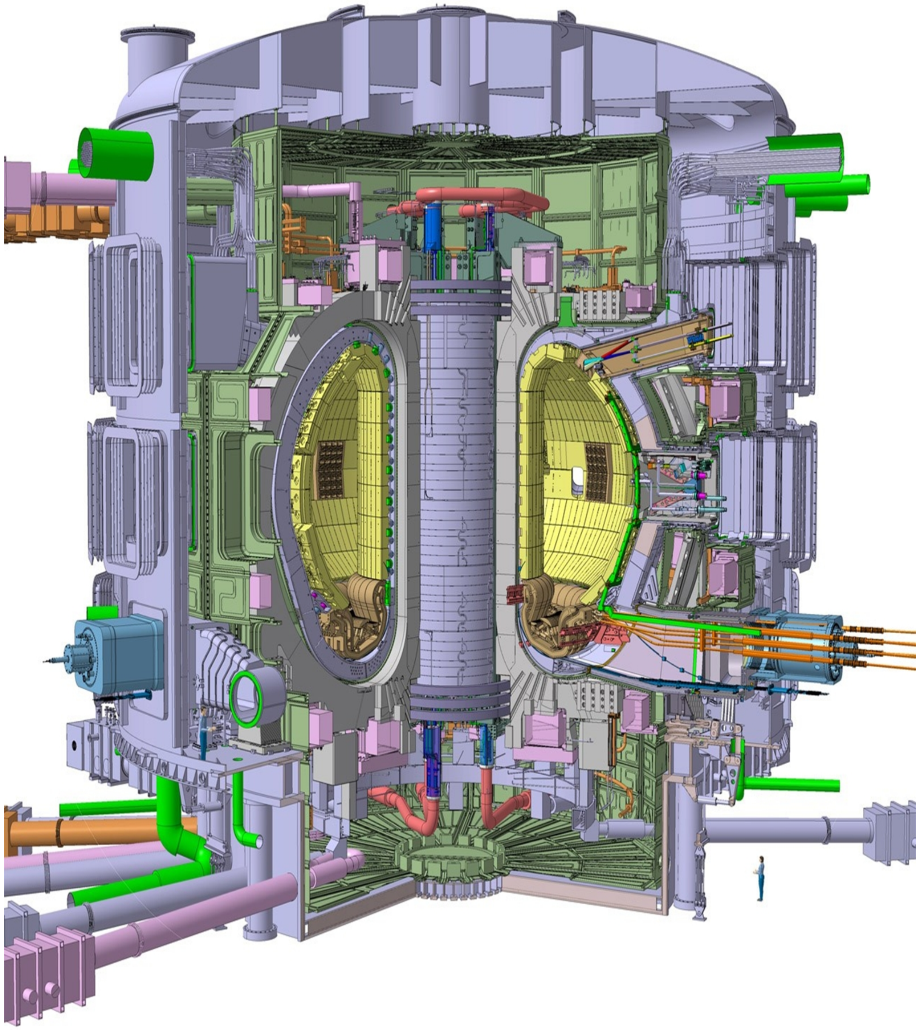
Problem Size: “Scaling Up"



DIII-D



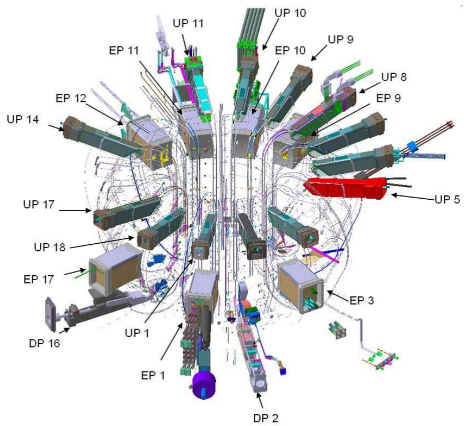
JET

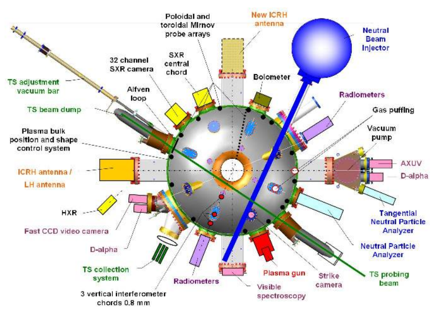


ITER

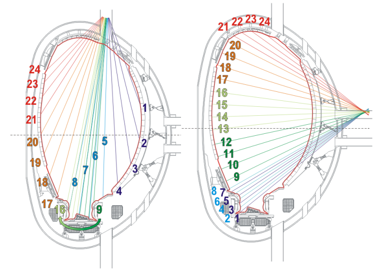
DIAGNOSTIC DATA SOURCES

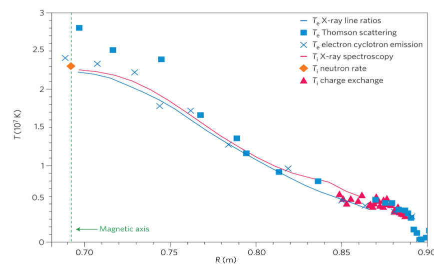
0D Scalar Data

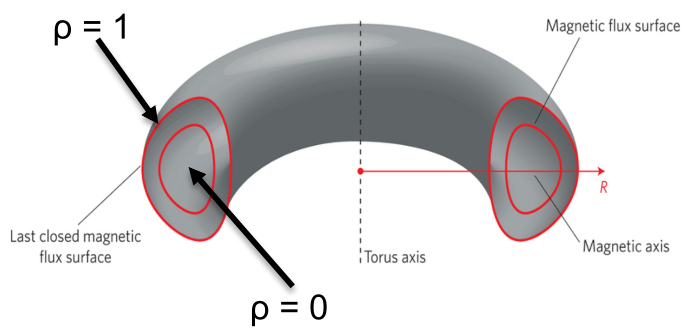




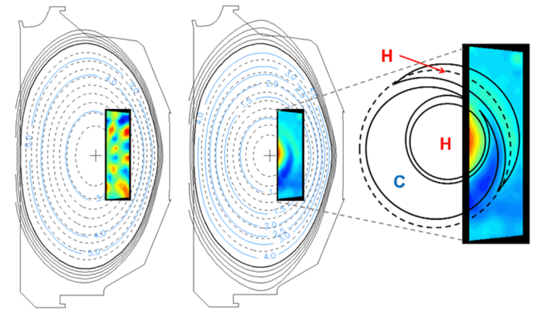
1D Profile Data



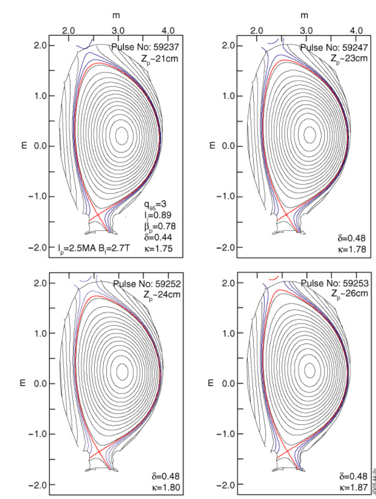




2D+ Data



ECEi imaging



Magnetic equilibria

EUROFUSION/JOIN-EUROPEAN TORUS (JET): Disruption Data

# Shots	Disruptive	Nondisruptive	Totals
Carbon Wall	324	4029	4353
Beryllium Wall (ILW)	185	1036	1221
Totals	509	5065	5574

JET produces ~ Terabyte (TB) of data per day

JET studies → 7 Signals of zero-D (scalar/zero-D) time traces, including	Data Size (GB)
Plasma Current	1.8
Mode Lock Amplitude	1.8
Plasma Density	7.8
Radiated Power	30.0
Total Input Power	3.0
d/dt Stored Diamagnetic Energy	2.9
Plasma Internal Inductance	3.0

~55 GB data collected from each JET shot

~ 350 TB total amount with multi-dimensional data currently being added

Machine Learning Workflow

Identify Signals
• Classifiers

Preprocessing and feature extraction

Normalization

Train model, Hyper parameter tuning

Use model for prediction

Princeton/PPPL DL predictions now advancing to *multi-D time trace signals (beyond zero-D)*

Measured sequential data arranged in patches of equal length for training

All data placed on appropriate numerical scale ~ O(1)
e.g., Data-based with all signals divided by their standard deviation

• All available data analyzed;
• Train LSTM (Long Short Term Memory Network) iteratively;
• Evaluate using ROC (Receiver Operating Characteristics) and cross-validation loss for every epoch (equivalent of entire data set for each iteration)

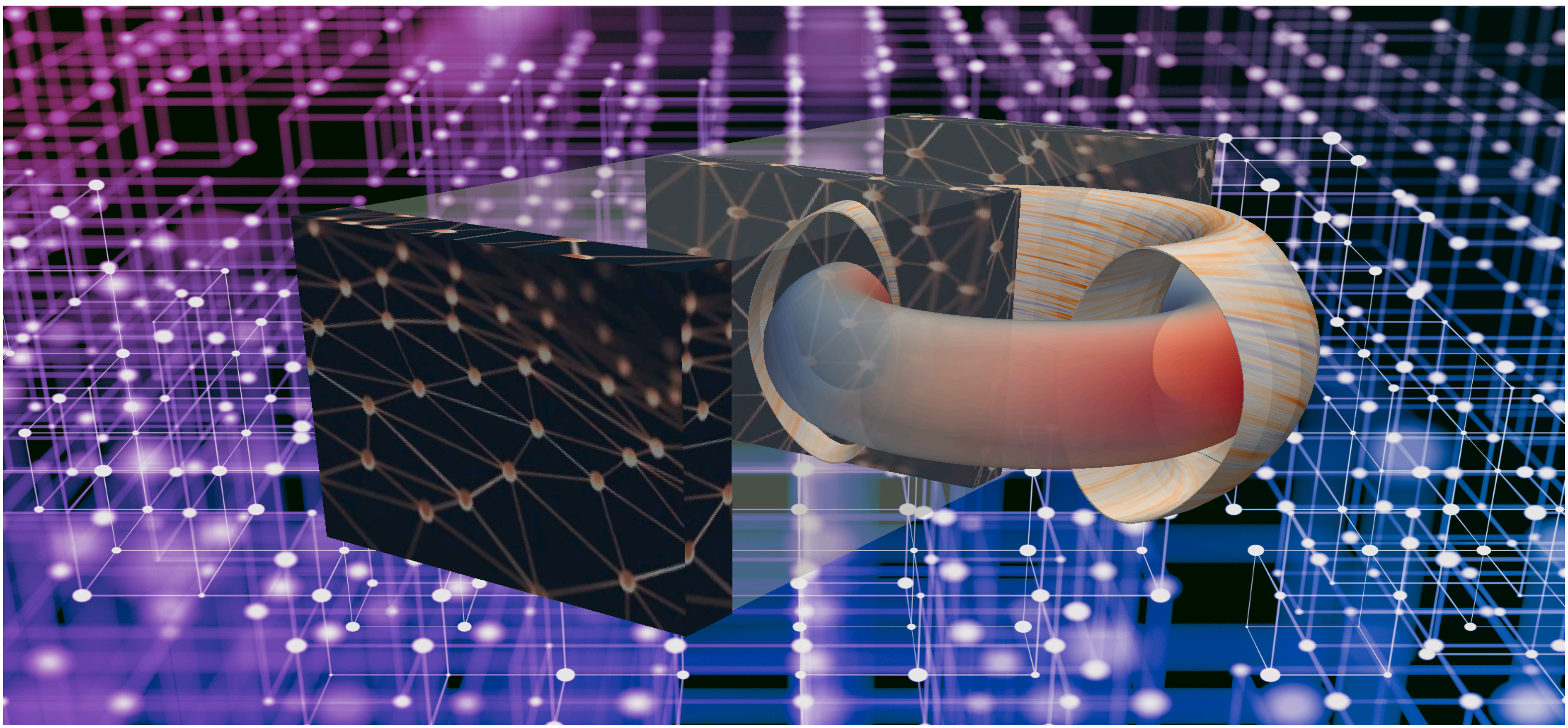
Apply ML/DL software on new data

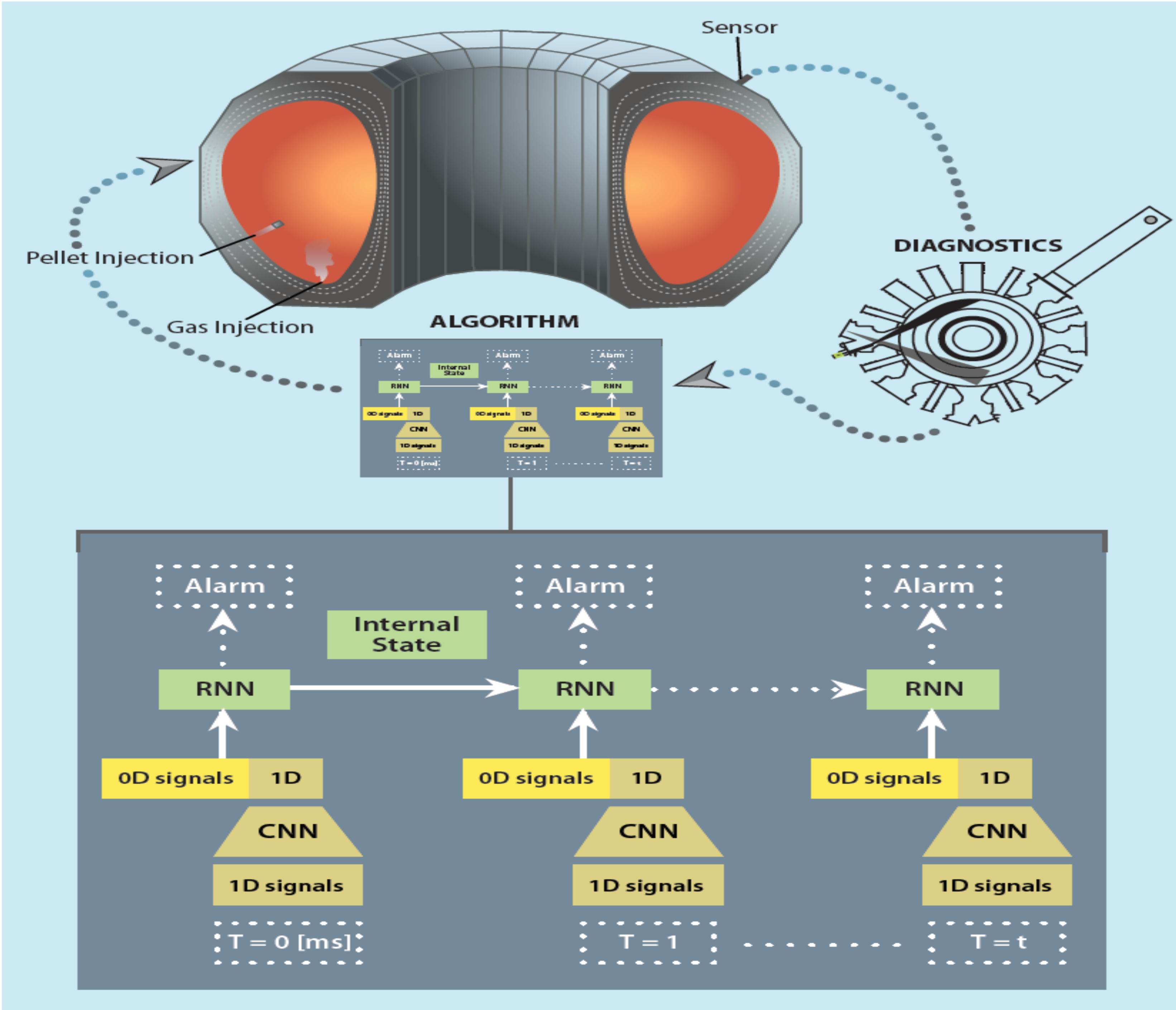
Artificial Intelligence/Deep Learning brings new technology to accelerate progress

"Predicting Disruptive Instabilities in Controlled Fusion Plasmas through Deep Learning"

NATURE: (accepted for publication, Jan. 2019, published, April 17, 2019 –
DOI: 10.1038/s41586-019-1116-4)

Princeton's Fusion Recurrent Neural Network code (FRNN) uses <u>*convolutional & recurrent neural network components*</u> *to integrate both spatial and temporal information for predicting disruptions in tokamak plasmas with* <u>*unprecedented accuracy and speed on top supercomputers*</u>

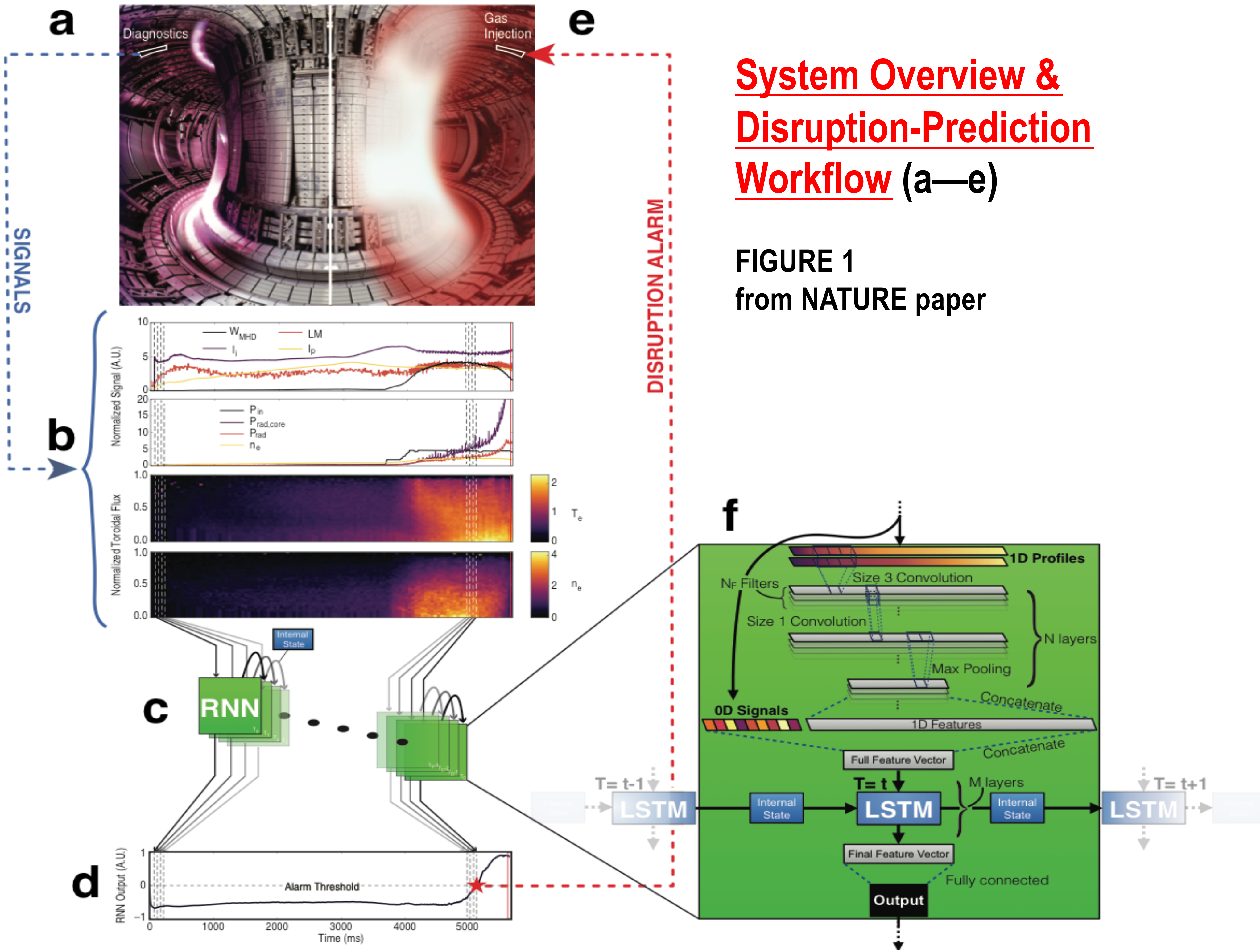




Sensor
Pellet Injection
Gas Injection
DIAGNOSTICS
ALGORITHM
Alarm
Alarm
Alarm
Internal State
RNN
RNN
RNN
0D signals
1D
0D signals
1D
0D signals
1D
CNN
CNN
CNN
1D signals
1D signals
1D signals
T = 0 [ms]
T = 1
T = t

System Overview & Disruption-Prediction Workflow (a—e)

FIGURE 1
from NATURE paper



System Overview & Disruption-Prediction Workflow (a—e)

Figure 1 Caption (from NATURE paper):

Top Image: interior view of the JET tokamak, with a non-disruptive plasma on the left and a disruptive plasma on the right.

Diagnostics **(a)** provide streams of sensory data **(b)** which are fed to the RNN-based deep learning algorithm **(c)** every 1 ms, producing a corresponding 'disruptivity' output at every time step **(d)**.

If output crosses a preset threshold value (dashed horizontal line), a disruption alarm is called (red star)
Alarm triggers mitigation action, such as gas injection **(e)** into the tokamak, to reduce the deleterious effects of the impending disruption **(f)**.

Detailed schematic of our deep-learning model:

Input data consist of scalar zero-dimensional (0D) signals and 1D profiles.

N layers of convolutional (containing NF filters each) and down-sampling (max-pooling) operations reduce dimensionality of the profile data and extract salient low-dimensional representations (features).

Features are concatenated with the 0D signals and fed into a multi-layer long-short term memory network (LSTM) with M layers, which also receives its internal state from the last time step as input.

The resulting final feature vector ideally contains salient information from the past temporal evolution and the present state of all signals.
This **vector** is fed through a fully connected layer to produce the output.

HIGHLIGHTS OF KEY ACHIEVEMENTS FEATURED IN NATURE PAPER (2019)

- ***Implementation of modern AI/Deep Learning advances enabled key achievements for Fusion Energy Science including:***

(1) Establishing ability to deal with one-dimensional physics signals for the first time – a significant improvement over previous Machine Learning R&D with focus on scalar-only “zero-D” signals.

(2) First demonstration of crucially-needed ability for predictive software trained on one experimental device (e.g., DIII-D tokamak) to make accurate predictions on another (e.g., the much larger, more powerful JET system) –> a key requirement for ITER relevance.

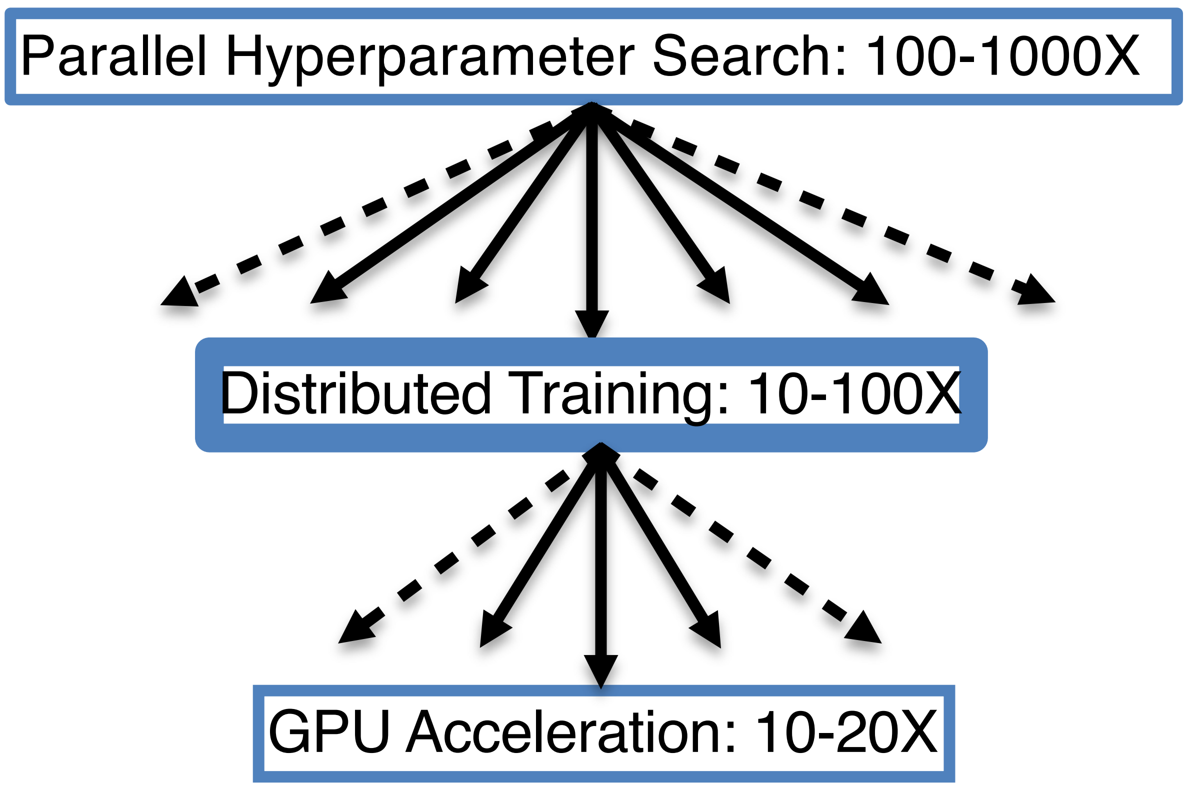
(3) Unique demonstration of AI/DL capability to efficiently utilize leadership class supercomputers to train predictive software-- e.g., Titan, Summit in US; Tsubame-3 in Japan, etc. – and exciting powerful systems in near future such as in (US) AURORA-21 (ALCF), Perlmutter (NERSC), Frontera (TACC), Frontier (OLCF) and in (JAPAN): ABCI & POST-K “FUGAKU”.

HPC Engagement

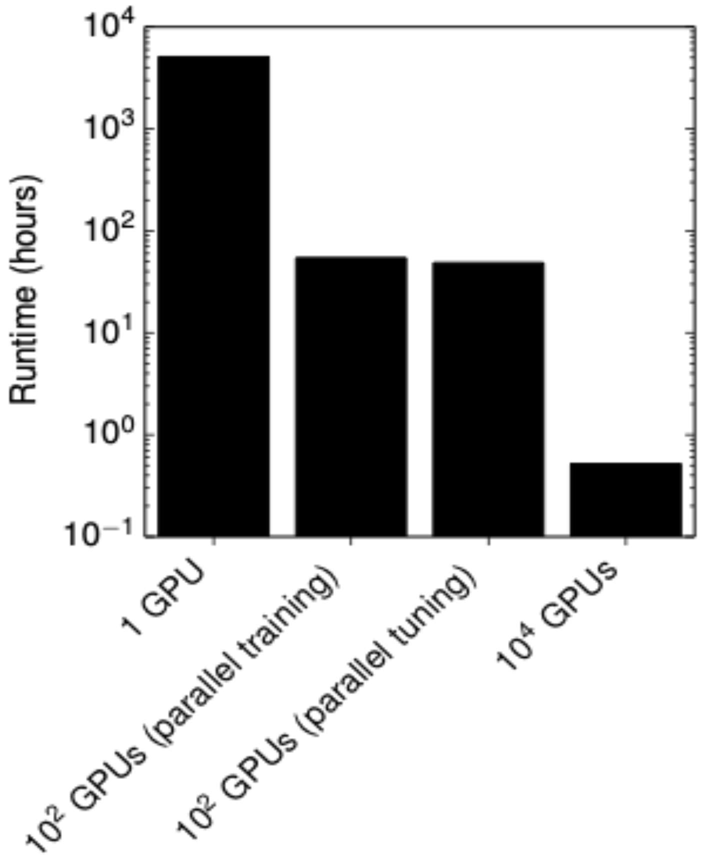
Challenges

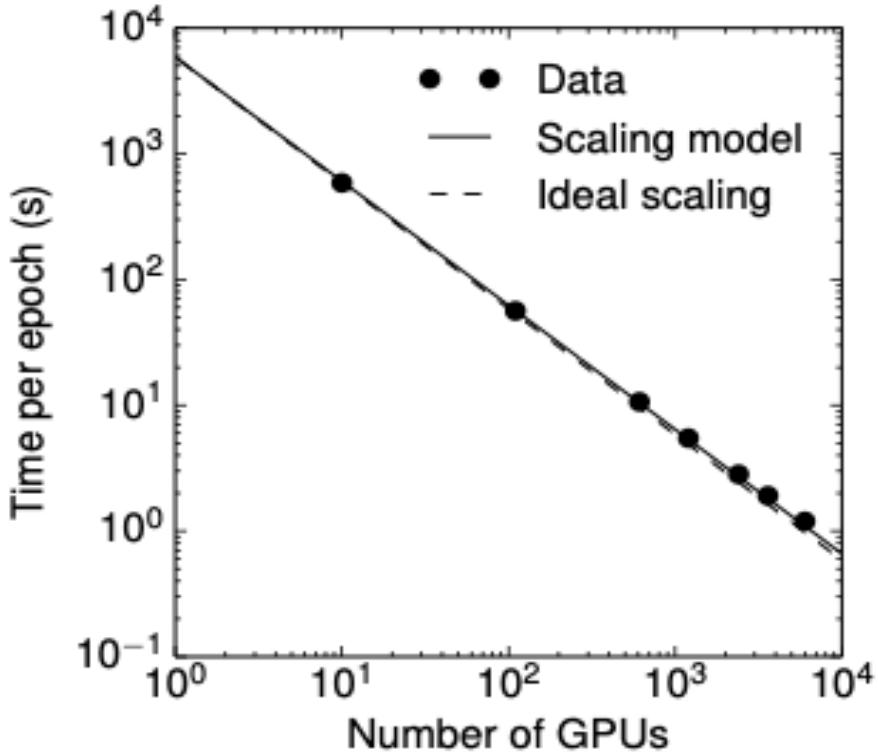
- Very large data sets (especially with multidimensional signals)
- Hyperparameter tuning (search high-D parameter space)

Levels of Parallelism



Distributed Training Scaling





FRNN DL/AI software reliably scaled to over 1K P-100 GPU's on TSUBAME 3.0 @ Tokyo Institute of Technology, Japan

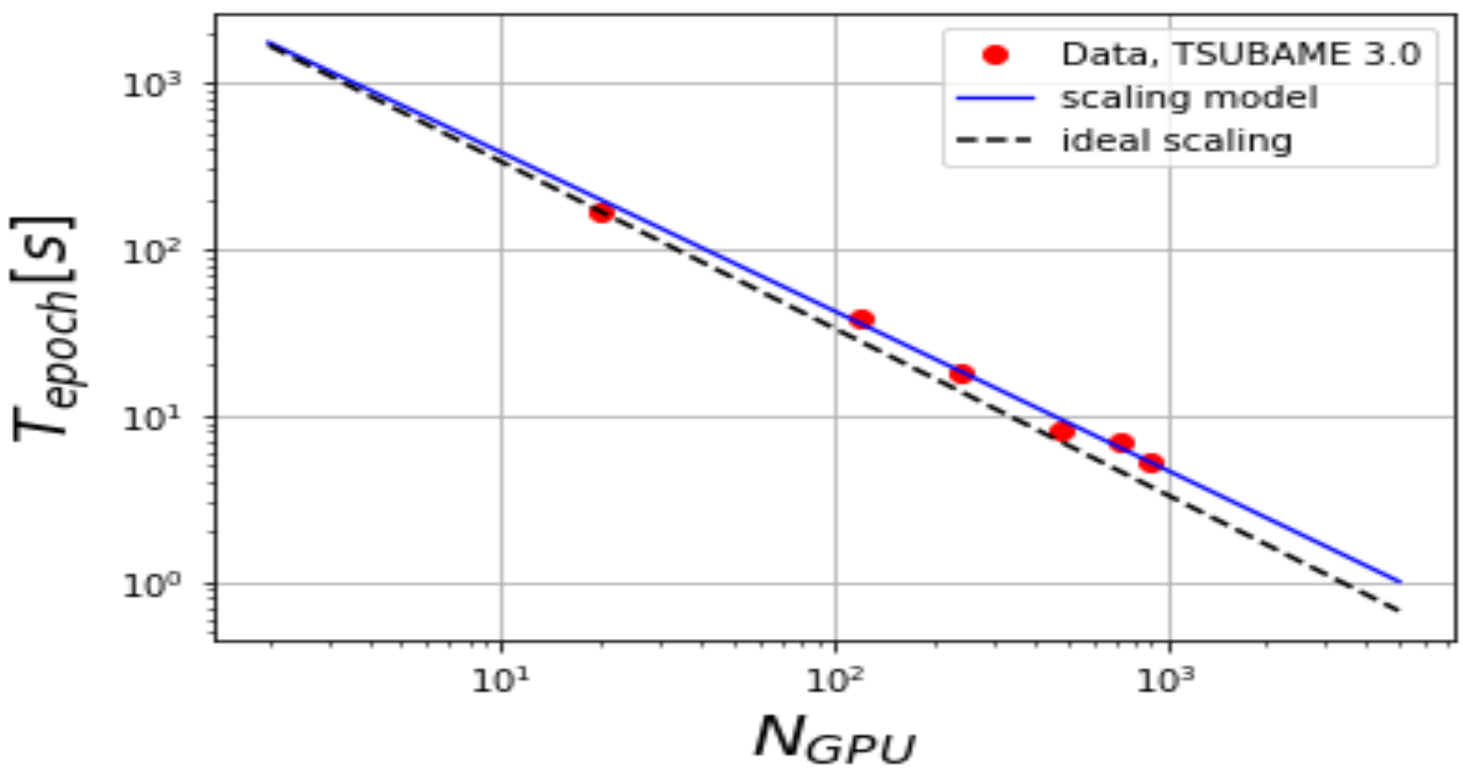
→ associated production runs contributed strongly to Hyper-pameter-Tuning-enabled physics advances !)

Illustrative results:

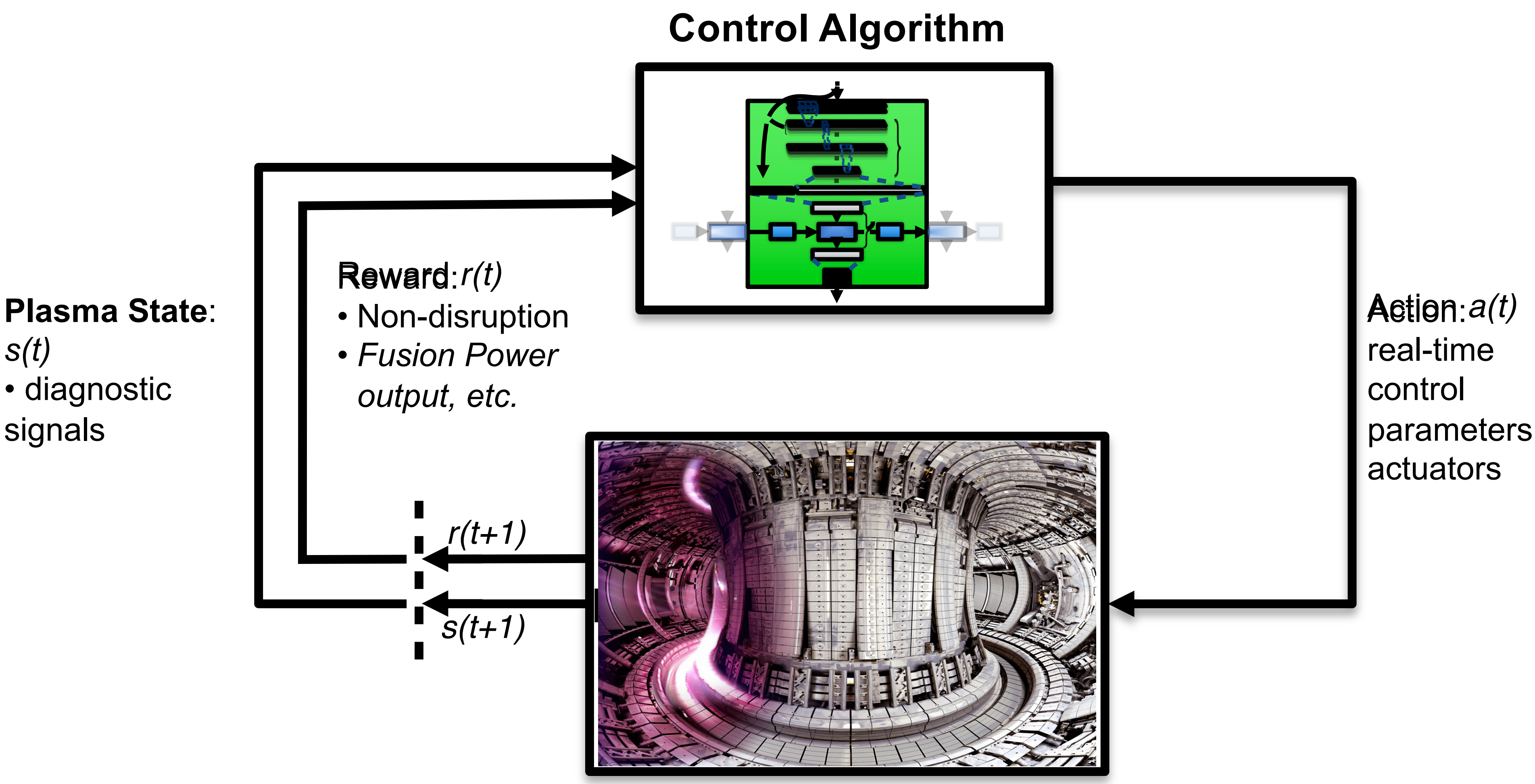
Tsubame 3.0 "Grand Challenge Runs" *(A. Svyatkovskii)*

–More than 1K Tesla P100 SXM2 GPUs, 4 GPUs per node, Nvlink communication

–Tensorflow+MPI, CUDA8, CuDNN 6



From Prediction to Control: Reinforcement Learning



SIGNIFICANT RECENT FRNN REAL-TIME ADVANCES

Integration of the AI/DL FRNN predictor into the DIII-D plasma control system at General Atomics, San Diego, CA has delivered encouraging progress in moving from prediction to real-time control:

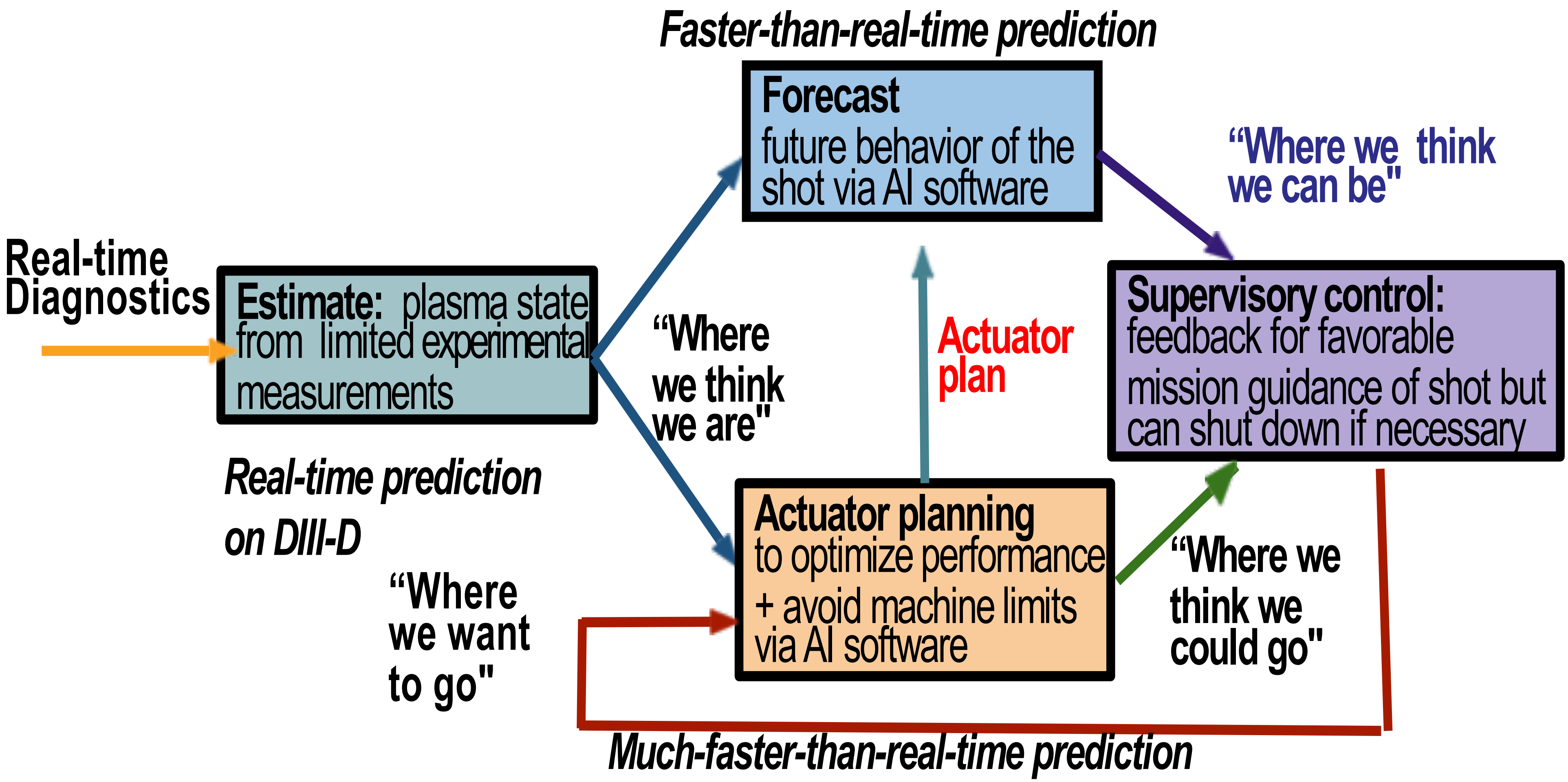
→ During D3D's "start-up" phase involving over 200 shots in May/June 2020, the FRNN inference engine shown to be readily functional during real-time operations;

→ Characteristic execution times of less than 1.7 msec. – were achieved in line with typical operating times for the PCS actuators;

→ Sets stage for systematic actuator testing targeting the demonstration of the ability to *modify the plasma state to avoid or delay the onset of disruptions.*

"Computing at the Edge": Real-Time Control of FES Plasma Disruption Experiments

(Requires Experimental/Advanced Diagnostic Expertise)



- Can we make our models fast & accurate enough?

--- e.g., via reinforcement learning/inference/ ......

- Can we make our models realistic enough?

--- e.g., via focused actuator planning with experimental partners

Ongoing Development of the FRNN
AI/Deep Learning Model Software Suite

FRNN [NATURE (April, 2019)]

- API (Application Prog. Interface): **Keras**
- Architecture: LSTM

Input

- 0D+1D data

Output

- Disruption score

FRNN [2020: In Progress R&D]

- API: **Keras** + Pytorch
- Architectures: LSTM; TCN; TT-LSTM

Input

- 0D+1D + 2D data

Output

- Disruption & real time sensitivity scores
- Physics-based signals in real-time application environment:
- Implement and test in DIII-D PCS

RAPID GROWTH & BROAD INVESTMENTS IN AI/DL/MACHINE LEARNING FOR THE FUTURE

Business World → Examples:

• **Amazon:** *Reformation of Amazon and other top businesses incorporating AI/DL/ML:*
https://www.wired.com/story/amazon-artificial-intelligence-flywheel/

• **Google Brain:** *AI/Deep Learning transformational impact in Natural Language Processing*
https://www.nytimes.com/2016/12/14/magazine/the-great-ai-awakening.html

R&D → Example: Cancer Research -- "Candle Project" with ECP

Exascale Computing Project (DOE & NIH) to identify optimal cancer treatment strategies, by building a ***scalable deep neural network code called the CANcer Distributed Learning Environment (CANDLE)***.

→development of predictive models for drug response, and automation of the analysis of information from millions of cancer patient records -- via developing, implementing, & testing DL/AI Algorithms and their benchmarks

• ***DL/AI advances in key application areas like Fusion Energy & Cancer Research can stimulate enhanced cross-disciplinary opportunties to leverage connections to enormous worldwide investments in AI/DL/ML R&D !***