The following are eight surrogates and a short description of each of them.

1- Predicting disruptive instabilities in controlled fusion plasmas through deep learning [1]

This paper presents a neural surrogate for forecasting disruptions in controlled fusion plasmas. The architecture combines both recurrent and convolutional components to extract spatiotemporal patterns from multimodal and high-dimensional sensory inputs. The surrogate successfully delivers predictions for reactors unseen during training; uses the information contained in high-dimensional diagnostic data, such as profiles, in addition to scalar signals; avoids the need for extensive feature engineering and selection; and enables rapid training times through high-performance computing

2- ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost [2]

This is a surrogate that can learn an accurate and transferable potential for organic molecules. It is trained with up to 8 heavy atoms in order to predict total energies for organic molecules containing four atom types: H, C, N, and O. ANI-1 is chemically accurate compared to reference DFT calculations on much larger molecular systems (up to 54 atoms) than those included in the training data set

3- CosmoGAN: creating high-fidelity weak lensing convergence maps using Generative Adversarial Networks [3]

This paper applies Generative Adversarial Networks(GAN) to the problem of generating weak lensing convergence maps. The GAN framework consists of a generator and a discriminator. The generator is trained to generate samples that aim to be indistinguishable from training data as judged by a competent discriminator. The discriminator is trained to judge whether a sample looks real or fake. The developed generator network produces maps that are described by, with high statistical confidence, the same summary statistics as the fully simulated maps.

4- Up to two billion times acceleration of scientific simulations with deep neural architecture search [4]

This work uses neural architecture search to build effective surrogates even with a limited number of training data. The method used successfully accelerates simulations by up to 2 billion times in 10 scientific cases including astrophysics, climate science, biogeochemistry, high energy density physics, fusion energy, and seismology, using the same super-architecture, algorithm, and hyperparameters. This approach also inherently provides emulator uncertainty estimation, adding further confidence in their use.

5- Machine Learning for Crystal Identification and Discovery [5]

This paper develops a surrogate for crystal identification and discovery. The surrogate can locate interesting structural regions of a complex phase diagram without prior information. In addition, the surrogate can automatically build order parameters that describe interesting structural behavior from data sets. The machine learning methods and structural descriptors are applicable anywhere that the local environment of a system needs to be characterized, even for complex crystals.

6- Hidden physics models: Machine learning of nonlinear partial differential equations [6]

This work proposes a framework that provides a general treatment of time-dependent and nonlinear partial differential equations, which can be of fundamentally different nature. The learning model that is used is Gaussian processes. The effectiveness of the proposed approach is demonstrated through a variety of canonical problems, spanning a number of scientific domains, including the Navier–Stokes, Schrödinger, Kuramoto–Sivashinsky, and time dependent linear fractional equations.

7- Machine learning surrogates for molecular dynamics simulations of soft materials [7]

This work develops a surrogate to replace molecular dynamics simulations of soft materials. The authors build a neural network-based regression model that successfully learns nearly all the interesting features associated with the output ionic density profiles over a broad range of ionic system parameters. The machine learning model generates predictions that are in excellent agreement with the results from molecular dynamics simulations. The inference time associated with the machine learning model is over a factor of 10,000 smaller than the corresponding parallel molecular dynamics simulation time.

8- Active learning of constitutive relation from mesoscopic dynamics for macroscopic modeling of non-Newtonian flows [8]

This paper develops a surrogate that is used alongside simulations to model non-newtenian fluids. The authors employ an active learning strategy based on Gaussian process regression (GPR) to minimize the number of expensive dissipative particle dynamics simulations, where adaptively selected dissipative particle dynamics simulations are performed only as necessary. Compared to traditional message-passing multiscale approaches, the active learning scheme used significantly increased the computational efficiency

use the covid susrrogates- add 1 to the table
1- The two billion surrogate - use it
2- come back to deep drive md

1- get the exact surrogate
2- get access to the code that is used to generate surrogate

Surrogate Number	Functionality Dimension	Problem Dimension	Algorithmic Dimension: Machine learning model	Algorithmic Dimension: Variability of the architecture
1	HPC simulations to train surrogates	Nuclear reactors	FRNN	fixed
2	replace simulation component	organic molecules	Fully-connected neural network with modefied HDNN method of Behler and Parrinello	fixed
3	replace a simulation	convergence maps	Generative Adversarial Network	fixed
4	wide range of possible functionalities	N/A	Convolutional Neural Network	Variable
5	replace part of a simulation	Crystal Identification	Artifical Neural Network (Not specified)	fixed
6	extract patterns	differential equations	Gaussian processes	fixed
7	replace a simulation	molecular dynamics	Artifical Neural Network (Not specified)	fixed
8	replace a simulation	fluid dynamics`	Gaussian process regression	fixed

Note that the following previously identified characteristics are not included in the table summary above. Those characteristics need to be removed from the classification or more clearly defined (except for surrogate architecture).

Characteristic	**Reason it is not included**
Performance improvement	1) all the explored surrogates outperformed “traditional methods” 2) definition of “traditional methods” is different from one problem to another
Generalizability	Definition is different from one surrogate to another
Link between ML and HPC	all the explored surrogates were HPCfor ML or HPC wasn’t used
Surrogate Architecture	1) The architecture of every surrogate is unique 2) Some architectures were too complex to describe in a table-format summary

Resources:

[1] Julian Kates-Harbeck, Alexey Svyatkovskiy, and William Tang. Predicting disruptive instabilities in controlled fusion plasmas through deep learning. Nature, 568(7753):526–531, 1 April 2019.

[2] Justin S Smith, Olexandr Isayev, and Adrian E Roitberg. Ani-1: an extensible neural network potential with dft accuracy at force field computational cost. Chemical science, 8(4):3192–3203, 2017.

[3] Mustafa Mustafa, Deborah Bard, Wahid Bhimji, Zarija Lukic, Rami Al-Rfou, and Jan M Kratochvil. ´ Cosmogan: creating high-fidelity weak lensing convergence maps using generative adversarial networks. Computational Astrophysics and Cosmology, 6(1):1–13, 2019.

[4] MF Kasim, D Watson-Parris, L Deaconu, S Oliver, P Hatfield, DH Froula, G Gregori, M Jarvis, S Khatiwala, J Korenaga, et al. Up to two billion times acceleration of scientific simulations with deep neural architecture search. arXiv preprint arXiv:2001.08055, 2020.

[5] Matthew Spellings and Sharon C Glotzer. Machine learning for crystal identification and discovery. AIChE journal. American Institute of Chemical Engineers, 64(6):2198–2206, 30 June 2018.

[6] Maziar Raissi and George Em Karniadakis. Hidden physics models: Machine learning of nonlinear partial differential equations. Journal of computational physics, 357:125–141, 15 March 2018.

[7] J C S Kadupitiya, Fanbo Sun, Geoffrey Fox, and Vikram Jadhao. Machine learning surrogates for molecular dynamics simulations of soft materials. Journal of computational science, 42:101107, 1 April 2020
[8] Lifei Zhao, Zhen Li, Bruce Caswell, Jie Ouyang, and George Em Karniadakis. Active learning of constitutive relation from mesoscopic dynamics for macroscopic modeling of non-newtonian flows. Journal of Computational Physics, 363:116–127, 2018.