AI and the Future of Theoretical Physics

Christian Ferko

Northeastern University and IAIFI, the
Institute for Artificial Intelligence and Fundamental Interactions

Splash 2025
Massachusetts Institute of Technology
March 15, 2025

“Attention is all you need.”

Much innovation in AI has been driven by neural networks and specifically **transformers** used in large language models (LLMs) like ChatGPT.

Remarkably, models like OpenAI’s Sora can now produce high-definition, realistic video from a simple text prompt. And reasoning models like o3-mini can now solve most **PhD-level** physics problems correctly.

Institute for AI and Fundamental Interactions.

AI is the most groundbreaking technological advance of our lifetimes, and it will impact all aspects of our life. But in today's talk, I would like to focus specifically on the interplay between **AI and physics**.

The NSF AI Institute for Artificial Intelligence and Fundamental Interactions (IAIFI)

Deep Learning (AI) + Deep Thinking (Physics) = Deeper Understanding

I currently work for **IAIFI**, which brings together researchers who are combining AI and physics in order to generate new insights in both fields.

Synergy in both directions.

The interaction between these two fields goes both ways: we can use **AI for physics**, but also **physics for AI**.

$$\begin{tikzcd}[column sep=huge] \textbf{Physics} \arrow[r, bend left, "\text{Apply physics techniques to understand models}"] & \textbf{AI} \arrow[l, bend left, "\text{Apply AI tools to solve problems}"] \end{tikzcd}$$

Roughly speaking, the main ideas of these two directions are:

1. It is hard to understand why "black box" models like neural networks work, but the conceptual framework of theoretical physics gives clues.
2. Many problems in theoretical physics are extremely difficult, so applying AI techniques gives us extra problem-solving power.

Roadmap.

Goal: talk about the back-and-forth between AI and physics, and lay out my vision for how AI will radically change the way that we do science.

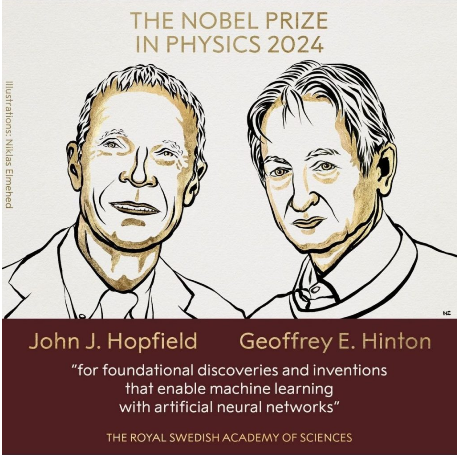
The plan is as follows:

- ☑ Part 1: Introduction and motivation.
- ☐ Part 2: Neural networks and quantum field theory (physics for AI).
- ☐ Part 3: Learning the geometry of extra dimensions (AI for physics)
- ☐ Part 4: 10,000 Einsteins and the future of science.

Part 2: Neural networks and quantum field theory (physics for AI).

Hopfield and Hinton.

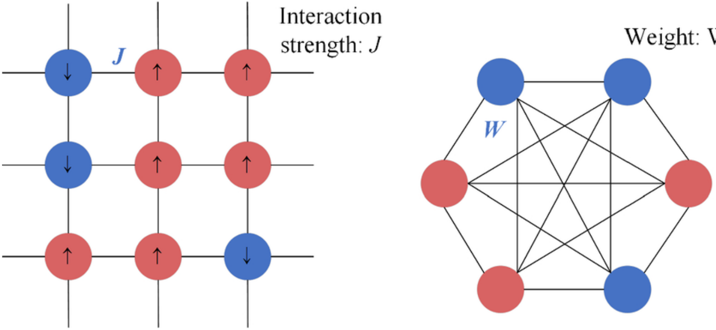
We begin our story with the 2024 Nobel Prize in Physics.



But why award the prize in **physics**? The reason is that the tools which inspired their work come from theoretical physics (statistical mechanics).

The Hopfield network and the Ising model.

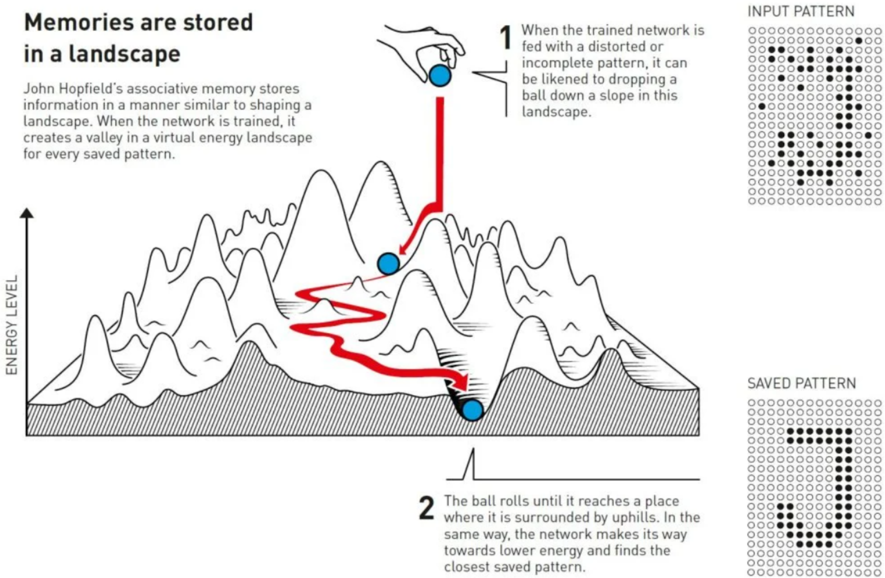
Hopfield's work – later extended by Hinton – is closely related to the *Ising model* in physics, which describes the behavior of electron spins in a magnetic material like iron. Spins can point "up" or "down".



(a) Ising Model (b) Hopfield Model

In the Ising model, spins prefer to align with their neighbors to minimize their energy. The strength of this preference is controlled by a number J.

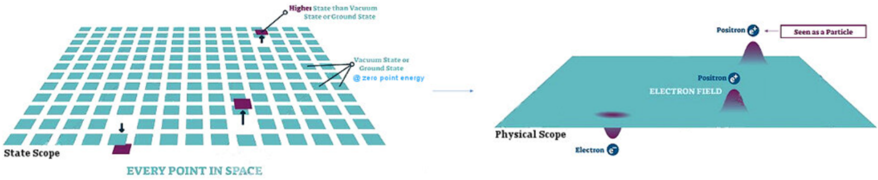
Exploring the energy landscape.



Connecting neurons and fields.

Hopfield and Hinton's work is the first example of a relationship between a *neural network* and a *field theory*.

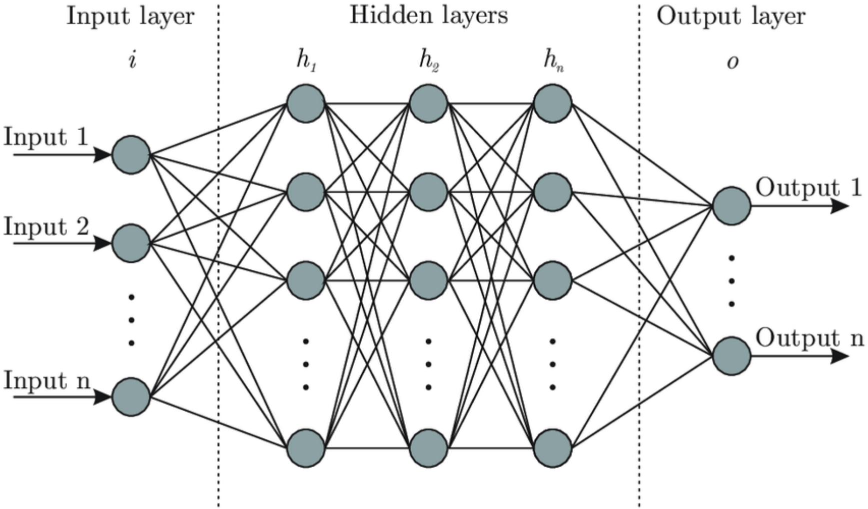
In physics, a field is a mathematical function that assigns a number to each point in space and time. That number can represent the spin of an atom, as in the Ising model, but there are many other examples.



In the standard model of particle physics, there is a value of the *electron field* at each point in space and time, another value for the *Higgs field*, etc.

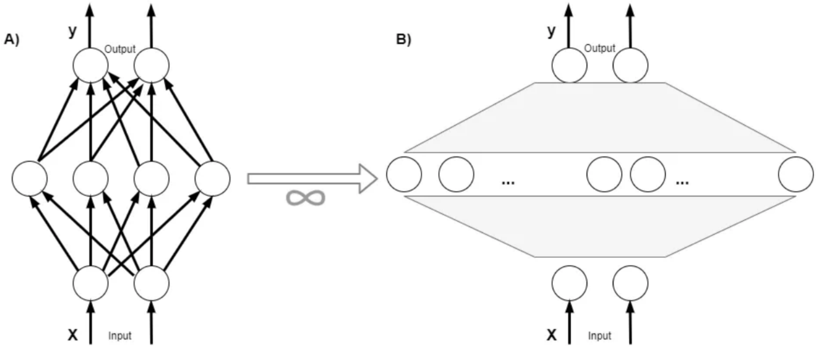
An algorithm inspired by the brain.

A *neural network* is a collection of nodes called *neurons* which are connected by *edges*. Here is a *feedforward* network with many layers:



The neural network - field theory correspondence.

In at least one example, a field theory (the Ising model) is related to a neural network (the Hopfield model). Are there other examples?



Yes! Surprise: as you make the "width" of *any* neural network larger and larger, it becomes *mathematically equivalent* to a free field theory.

Using field theory to understand AI.

Physicists have developed an *enormous* set of tools for studying field theory. The NN-FT correspondence means that we can also apply this toolkit to understand the behavior of neural networks!

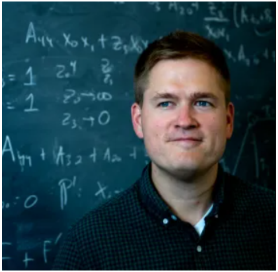
Unitarity in Neural Network Field Theories

Christian Ferko[a,b] **and James Halverson**[a,b]

[a] *Department of Physics, Northeastern University, Boston, MA 02115, USA*

[b] *The NSF Institute for Artificial Intelligence and Fundamental Interactions*

c.ferko@northeastern.edu, j.halverson@northeastern.edu

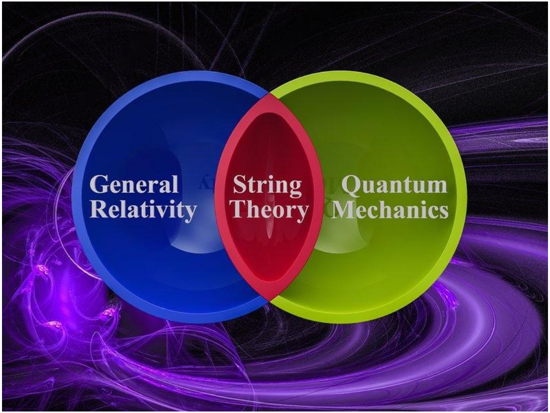


This is an example of **physics for AI**. Work by IAIFI researchers, including Jim and myself, is ongoing in this exciting area.

Part 3: Learning the geometry of extra dimensions (AI for physics).

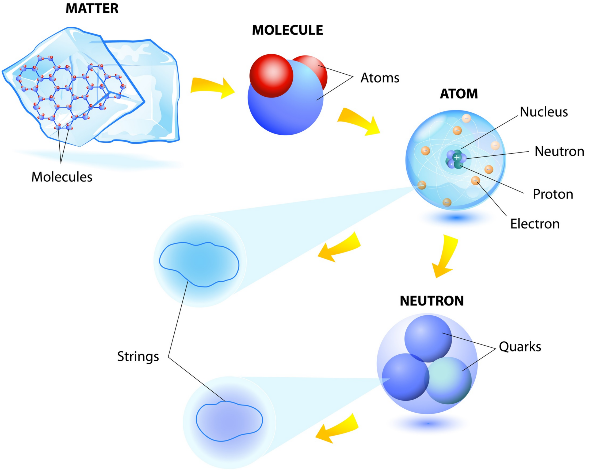
Quantum gravity.

In the last 50 years, we learned how to combine **general relativity** with the principles of **quantum mechanics** using the framework of *string theory*.



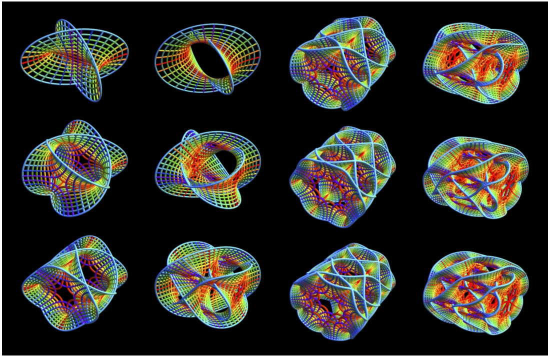
Strings as fundamental constituents.

Within string theory, all of the fundamental particles that we learn about – electrons, the quarks that make up protons, and even the photons which are particles of light – are simply strings vibrating in different patterns.



Extra dimensions.

For mathematical consistency, (super)string theory requires *ten* spacetime dimensions, but we seem to observe only four (3 space, 1 time).



The remaining 6 dimensions must be curled up into a tiny space, small enough to evade detection. We often choose this space to be a *Calabi-Yau* manifold (pictured above) due to their convenient theoretical properties.

The geometry of Calabi-Yaus.

To make **physical predictions** about string theory with a particular shape for the extra dimensions, we need some mathematical information about the geometry of the curled-up space (namely the *metric*).

For many years, this was impossible! We could prove the *existence* of a metric on a Calabi-Yau, but didn't know how to find it.

Then a colleague of mine at Northeastern, Fabian Ruehle, applied AI to this problem.

Machine learning Calabi-Yau metrics.

With collaborators, Fabian trained a neural network to *learn* the metric on a CY manifold. They did this by asking the network to *minimize* certain quantities which need to be zero for mathematical consistency of the CY.

Moduli-dependent Calabi-Yau and SU(3)-structure metrics from Machine Learning

Lara B. Anderson, Mathis Gerdes, James Gray, Sven Krippendorf, Nikhil Raghuram, Fabian Ruehle

We use machine learning to approximate Calabi-Yau and SU(3)-structure metrics, including for the first time complex structure moduli dependence. Our new methods furthermore improve existing numerical approximations in terms of accuracy and speed. Knowing these metrics has numerous applications, ranging from computations of crucial aspects of the effective field theory of string compactifications such as the canonical normalizations for Yukawa couplings, and the massive string spectrum which plays a crucial role in swampland conjectures, to mirror symmetry and the SYZ conjecture. In the case of SU(3) structure, our machine learning approach allows us to engineer metrics with certain torsion properties. Our methods are demonstrated for Calabi-Yau and SU(3)-structure manifolds based on a one-parameter family of quintic hypersurfaces in $\mathbb{P}^4$.

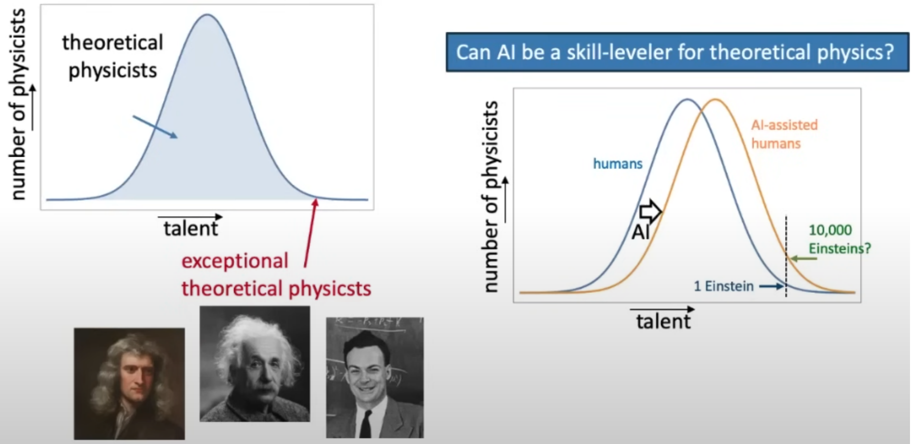
This is an example of **AI for physics** and is a major breakthrough!

For the first time, we can use *concrete numerical data* to make specific predictions about string theory given a choice for the extra dimensions.

Part 4: 10,000 Einsteins and the future of science.

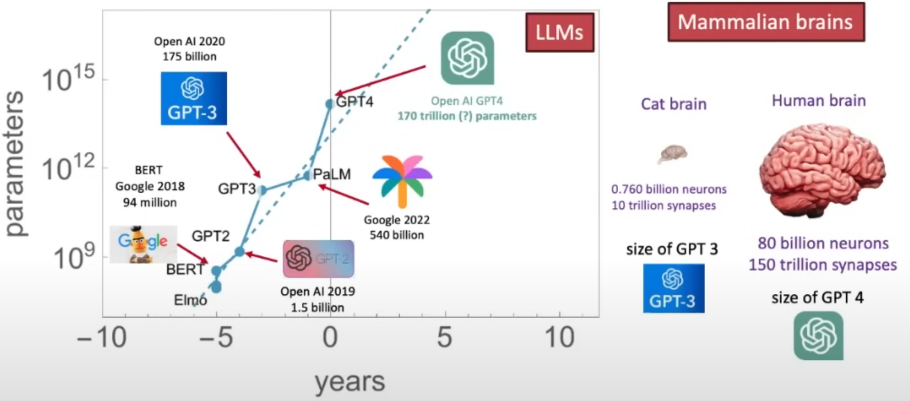
Leveling up our physics abilities.

We have seen that AI for physics is a “power-up” that allows us to tackle problems that would otherwise be too difficult.



Beyond the short-term.

In the next few years, AI-assisted science might make *all* researchers more powerful, leading to 10,000 Einsteins. But what about in the long term?



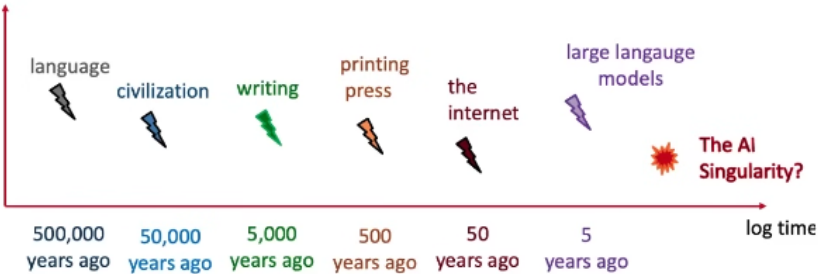
LLMs have been growing about 10 times more powerful each year, and GPT4 probably has more parameters than the human brain.

Recursive self-improvement.

One idea – attention – was the spark that ignited recent progress in LLMs.

Could AI discover the *next* big idea that leads to massive progress in AI?

It seems plausible – GPT4 (or 4.5) is already more creative than, and can code better, than most humans. GPT5 will be much better.



Once this happens, we reach the point where AI can improve itself.

Can humans ever understand the theory of everything?

It might be that we can *never* fully understand the implications of string theory because we are limited by our brains. After all, no amount of careful instruction and practice would ever succeed in teaching a cat to play chess.

But even so, a super-intelligent self-improving AI might be able to understand it. In the far future, science may consist of attempting to learn what we can from an AI that understands reality better than we ever will.

Conclusion.

We live in the dawn of the AI era.

We are fortunate to be alive **at the exact crossover point** where the complexity of AI models has just exceeded that of the human brain.

As we continue to push for more powerful AI models, we can expect:

1. the intellectual toolkit of the theoretical physicist will give us deeper insights into how and why AI works;
2. the awe-inspiring power of AI models will allow us to make serious progress in understanding the fundamental nature of reality; and
3. once AI learns to improve itself, we will see models that can perform science research on their own and explain the results to us.

We may even live to see the end of human-driven science. But whether this is a good thing or a bad thing is a topic for another talk.

Thank you for your attention!

Email cferko@alum.mit.edu with questions.