



WILL WE BE REPLACED BY AI?

JIA LIU

HITOSHI-FEST, DECEMBER 16TH, 2024

サイカル journal
SCIENCE & CULTURE by NHK

2024年に大注目の科学は？
物理学者 村山齊さんに聞
— NHK 2024.01.13



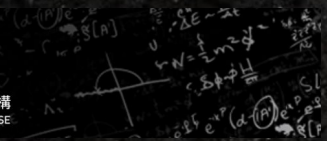
Tea time, Kavli IPMU



Lawrence Berkeley National Laboratory
Berkeley, California, September 28, 2023

KAVLI
IPMU

東京大学 国際高等研究所 カブリ数物連携宇宙研究機構
KAVLI INSTITUTE FOR THE PHYSICS AND MATHEMATICS OF THE UNIVERSE



Inflation & Early Universe

Phenomenology

Quantum Field Theory

String Theory

Mathematics

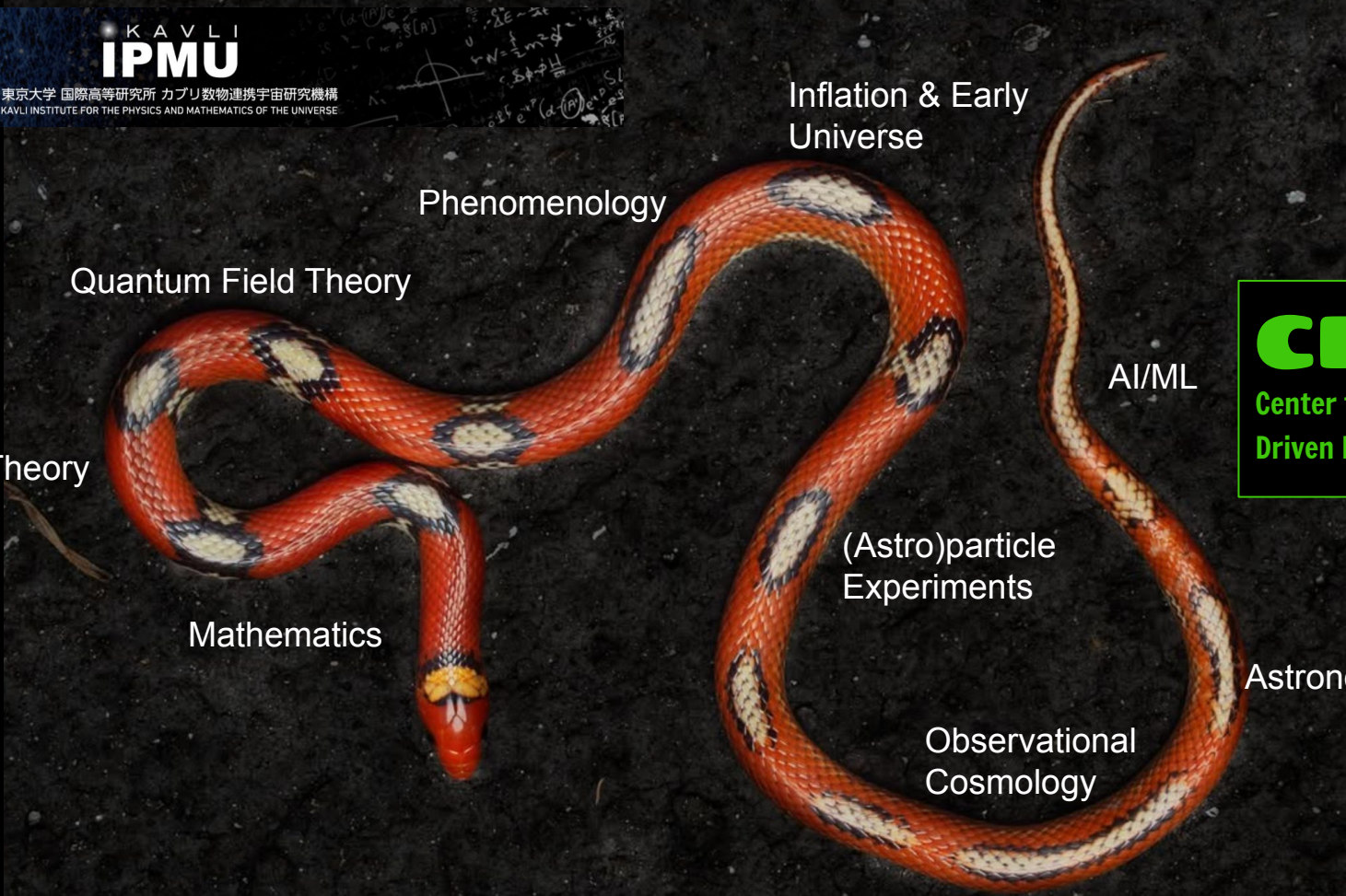
AI/ML

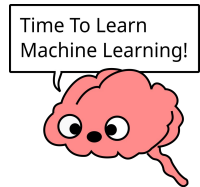
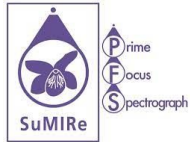
CD3
Center for Data-Driven Discovery

(Astro)particle Experiments

Astronomy

Observational Cosmology





Physics, cosmology, social science projects at **CD3**

CD3 is 1 year old now



Future Science with CMB x LSS (April 10-14, 2023, YITP, Kyoto)



Astro AI with Fugaku (Sept. 11-12, 2023, U Tsukuba)



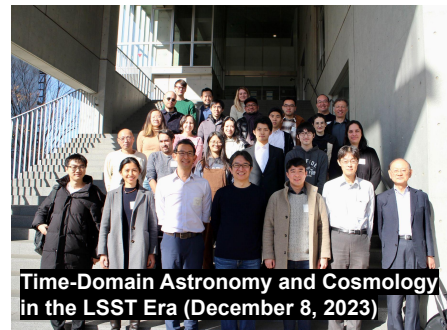
Year 1 of CD3: building global interdisciplinary connection



AI-driven discovery in physics and astrophysics (January 22-26, 2024)



CD3 Opening Symposium (April 19-20, 2023)



Time-Domain Astronomy and Cosmology in the LSST Era (December 8, 2023)

Dark matter detection at COSI (March 21-22, 2024)



Baryons in the Universe 2024 (April 8-12, 2024)



A³ Net AstroAI Asian Network
cd3.ipmu.jp/a3n

Upcoming Summer School:

August 18-22, 2025, KIAS, South Korea

Center for Data-Driven Discovery (CD3), Kavli IPMU
Theoretical Joint Research (TJR) Project, Osaka University
Fudan University
University of Hawaii
University of Hong Kong
Kavli Institute for Astronomy and Astrophysics (KIAA), Peking University
Korea Institute for Advanced Study (KIAS)
Kobayashi-Maskawa Institute for the Origin of Particles and the Universe (KMI), Nagoya University
Leung Center for Cosmology and Particle Astrophysics
National Taiwan University
Shanghai Jiaotong University
Shanghai Astronomical Observatory, Chinese Academy of Sciences
Tsinghua University
Program for Fugaku: JPMXP1020230406, University of Tsukuba

Year 2 of CD3: building AI x Physics educational program in Asia



First Summer School: September 2-6, 2024, Osaka, Japan

Future Science with CMB x LSS (April 10-14, 2023, YITP, Kyoto)



Baryons in the Universe 2024 (April 8-12, 2024)



AI-driven discovery in physics and astrophysics (January 22-26, 2024)



CD3 x CMB x Astro Seminar by Jo Dunkley (Feb. 5, 2024)



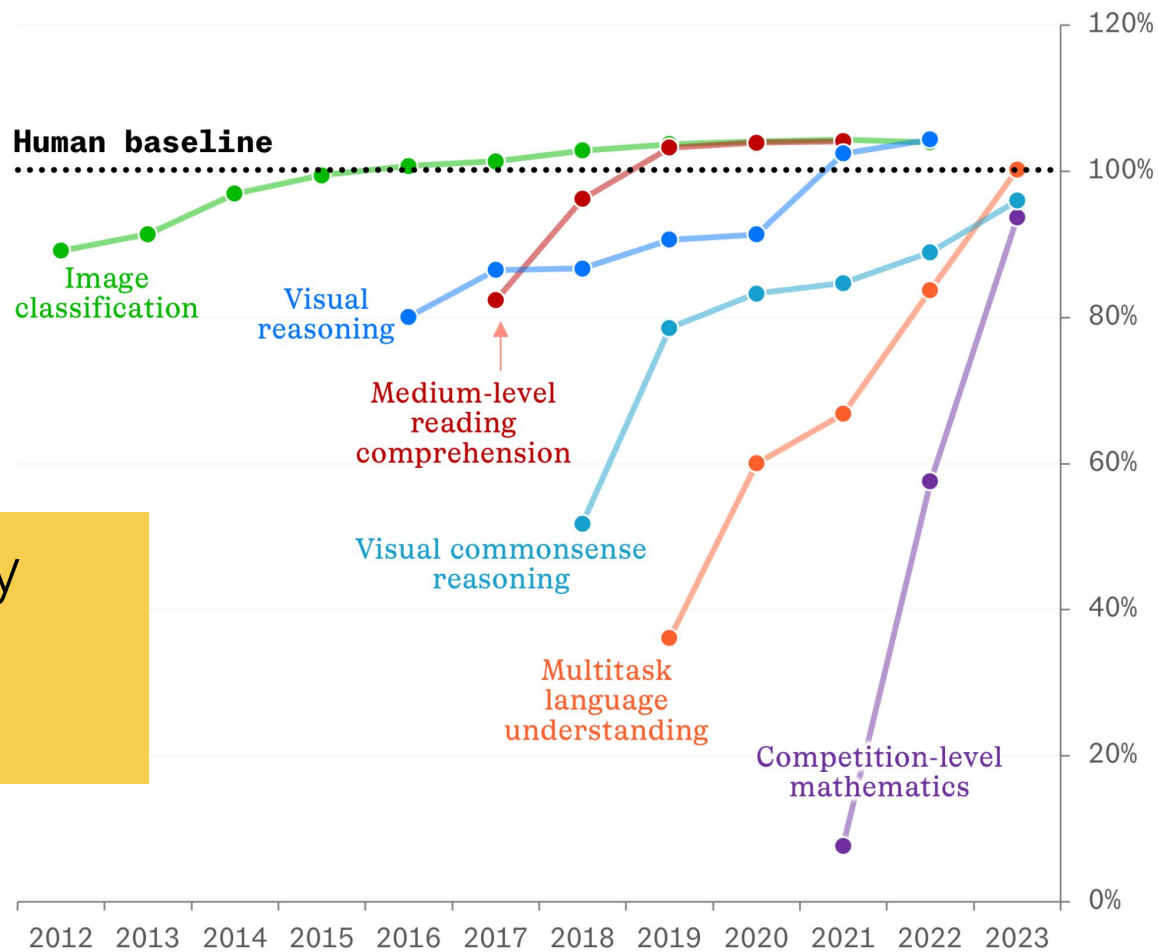
Women & minority lunch at CD3

A close-up shot of the character WALL-E from the Pixar movie 'WALL-E'. He is holding a Rubik's cube in his right hand. The background is a dark, industrial setting with some lights. The text 'WILL WE BE REPLACED BY AI?' is overlaid in yellow, bold, sans-serif font across the middle of the image. On WALL-E's chest, there is a panel with a sun icon and the text 'SOLAR CHARGE LEVEL' and 'WALL-E' is visible on his lower body.

WILL WE BE REPLACED BY AI?



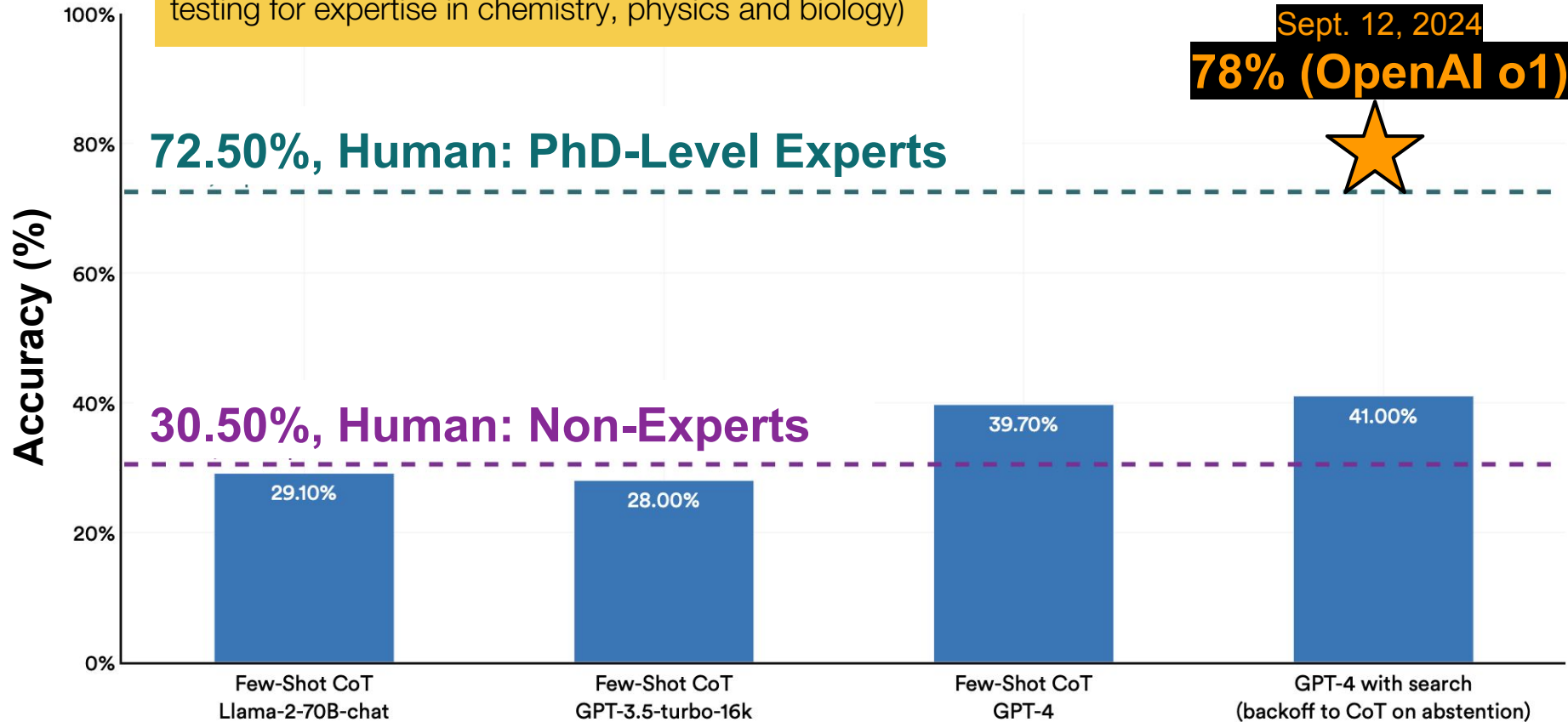
1997 **Deep Blue** defeated world chess champion **Garry Kasparov** in a six-game match



AI now performs many intellectual tasks at a **superhuman** level

GPQA Benchmark

(Google-Proof Q&A, a difficult intelligence benchmark testing for expertise in chemistry, physics and biology)



GPQA Benchmark

(Google-Proof Q&A, a difficult intelligence benchmark testing for expertise in chemistry, physics and biology)

Astrophysics

Astronomers are studying a star with a Teff of approximately 6000 K. They are interested in spectroscopically determining the surface gravity of the star using spectral lines (EW < 100 mÅ) of two chemical elements, E11 and E12. Given the atmospheric temperature of the star, E11 is mostly in the neutral phase, while E12 is mostly ionized. Which lines are the most sensitive to surface gravity for the astronomers to consider?

- A) E12 I (neutral)
 - B) E11 II (singly ionized)
 - C) E12 II (singly ionized)
 - D) E11 I (neutral)
-

Quantum Mechanics

Suppose we have a depolarizing channel operation given by $E(\rho)$. The probability, p , of the depolarization state represents the strength of the noise. If the Kraus operators of the given state are $A_0 = \sqrt{1 - \frac{3p}{4}}$, $A_1 = \sqrt{\frac{p}{4}}X$, $A_2 = \sqrt{\frac{p}{4}}Y$, and

$A_3 = \sqrt{\frac{p}{4}}Z$. What could be the correct Kraus Representation of the state $E(\rho)$?

- A) $E(\rho) = (1 - p)\rho + \frac{p}{3}X\rho X + \frac{p}{3}Y\rho Y + \frac{p}{3}Z\rho Z$
 - B) $E(\rho) = (1 - p)\rho + \frac{p}{3}X\rho^2 X + \frac{p}{3}Y\rho^2 Y + \frac{p}{3}Z\rho^2 Z$
 - C) $E(\rho) = (1 - p)\rho + \frac{p}{4}X\rho X + \frac{p}{4}Y\rho Y + \frac{p}{4}Z\rho Z$
 - D) $E(\rho) = (1 - p)\rho^2 + \frac{p}{3}X\rho^2 X + \frac{p}{3}Y\rho^2 Y + \frac{p}{3}Z\rho^2 Z$
-



Attention? Attention!

Attention Is All You Need

Ashish Vaswani*
Google Brain
avaswani@google.com

Noam Shazeer*
Google Brain
noam@google.com

Niki Parmar*
Google Research
nikip@google.com

Jakob Uszkoreit*
Google Research
usz@google.com

Llion Jones*
Google Research
llion@google.com

Aidan N. Gomez* †
University of Toronto
aidan@cs.toronto.edu

Łukasz Kaiser*
Google Brain
lukaszkaizer@google.com

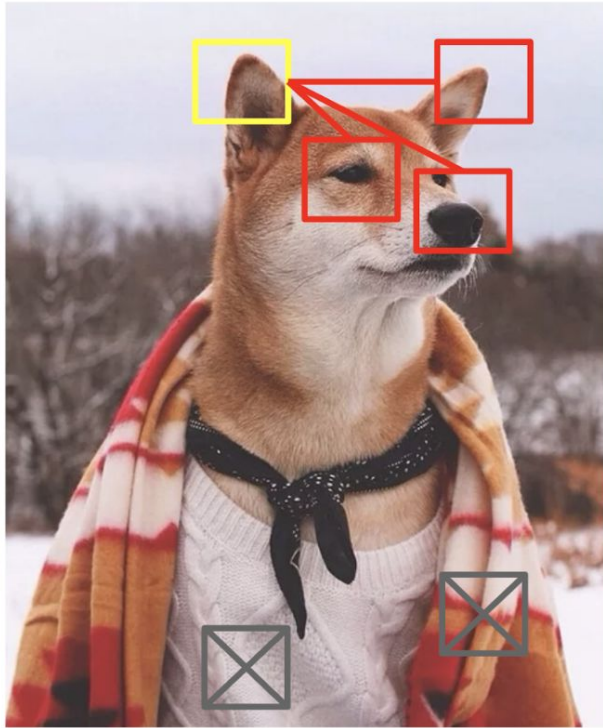
Illia Polosukhin* ‡
illia.polosukhin@gmail.com

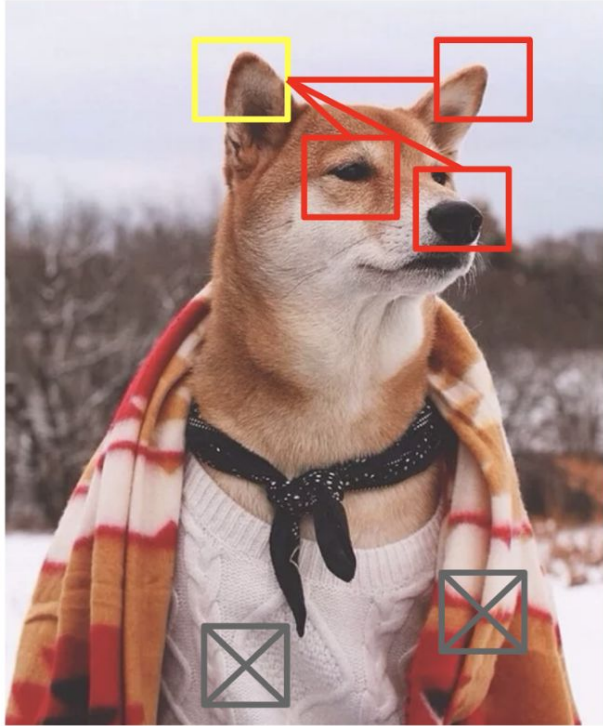
Abstract

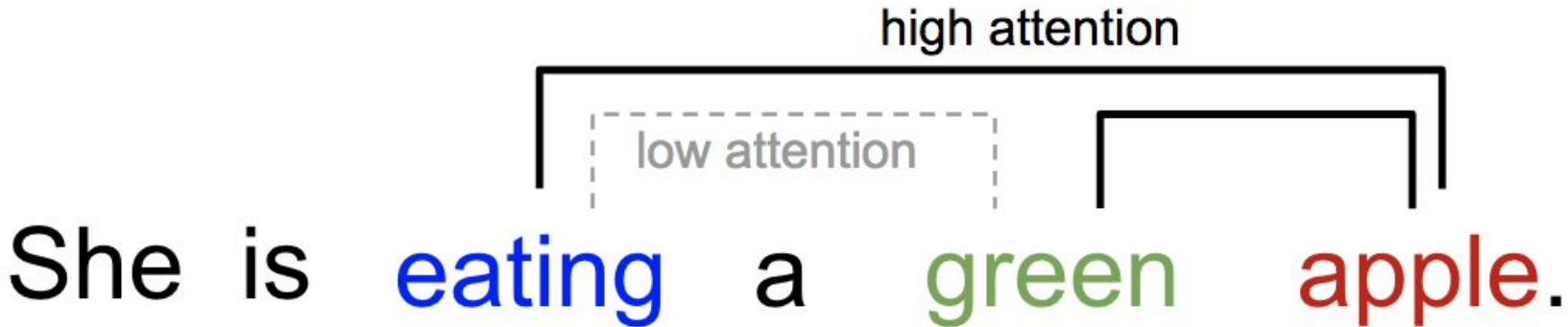
The dominant sequence transduction models are based on complex recurrent or convolutional neural networks that include an encoder and a decoder. The best performing models also connect the encoder and decoder through an attention mechanism. We propose a new simple network architecture, the Transformer, based solely on attention mechanisms, dispensing with recurrence and convolutions entirely. Experiments on two machine translation tasks show these models to be superior in quality while being more parallelizable and requiring significantly less time to train. Our model achieves 28.4 BLEU on the WMT 2014 English-to-German translation task, improving over the existing best results, including ensembles, by over 2 BLEU. On the WMT 2014 English-to-French translation task, our model establishes a new single-model state-of-the-art BLEU score of 41.8 after training for 3.5 days on eight GPUs, a small fraction of the training costs of the best models from the literature. We show that the Transformer generalizes well to other tasks by applying it successfully to English constituency parsing both with large and limited training data.











Generative Agents: Interactive Simulacra of Human Behavior

Joon Sung Park
Stanford University
Stanford, USA
joonspk@stanford.edu

Joseph C. O'Brien
Stanford University
Stanford, USA
jobrien3@stanford.edu

Carrie J. Cai
Google Research
Mountain View, CA, USA
cjcai@google.com

Meredith Ringel Morris
Google DeepMind
Seattle, WA, USA
merrie@google.com

Percy Liang
Stanford University
Stanford, USA
pliang@cs.stanford.edu

Michael S. Bernstein
Stanford University
Stanford, USA
msb@cs.stanford.edu



A community of 25
unique agents inhabits
Smallville...

Figure 1: Generative agents are believable simulacra of human behavior for interactive applications. In this work, we demonstrate generative agents by populating a sandbox environment, reminiscent of The Sims, with twenty-five agents. Users can observe and intervene as agents plan their days, share news, form relationships, and coordinate group activities.



A Valentine's Day party at Hobbs Cafe



Isabella Rodriguez, at Hobbs Cafe, is initialized with an intent to plan a Valentine's Day party from 5 to 7 p.m. on February 14th. From this seed, the agent proceeds to invite friends and customers when she sees them at Hobbs Cafe or elsewhere. Isabella then spends the afternoon of the 13th decorating the cafe for the occasion. Maria, a frequent customer and close friend of Isabella's, arrives at the cafe. Isabella asks for Maria's help in decorating for the party, and Maria agrees. Maria's character description mentions that she has a crush on Klaus. That night, Maria invites Klaus, her secret crush, to join her at the party, and he gladly accepts. On Valentine's Day, five agents, including Klaus and Maria, show up at Hobbs Cafe at 5 pm, and they enjoy the festivities.

A Valentine's Day party at Hobbs Cafe



Isabella Rodriguez, at Hobbs Cafe, is initialized with an intent to plan a Valentine's Day party from 5 to 7 p.m. on February 14th. From this seed, the agent proceeds to invite friends and customers when she sees them at Hobbs Cafe or elsewhere. Isabella then spends the afternoon of the 13th decorating the cafe for the occasion. Maria, a frequent customer and close friend of Isabella's, arrives at the cafe. Isabella asks for Maria's help in decorating for the party, and Maria agrees. Maria's character description mentions that she has a crush on Klaus. That night, Maria invites Klaus, her secret crush, to join her at the party, and he gladly accepts. On Valentine's Day, five agents, including Klaus and Maria, show up at Hobbs Cafe at 5 pm, and they enjoy the festivities.

A young woman with long, wavy blonde hair and light blue eyes is shown in a close-up, three-quarter view. She is looking off to the right with a thoughtful expression. The background is dark and out of focus, suggesting an indoor setting with some equipment. Overlaid on the image is the text "WILL WE BE REPLACED BY AI?" in a bold, yellow, sans-serif font.

WILL WE BE REPLACED BY AI?

Soundbytes

Game commentary in snack-sized mouthfuls



"I want AI to do my laundry and dishes so that I can do art and writing, not for AI to do my art and writing so that I can do my laundry and dishes."

Author and videogame enthusiast **Joanna Maciejewska** nails it (although bathroom cleaning goes ahead of laundry and dishes)

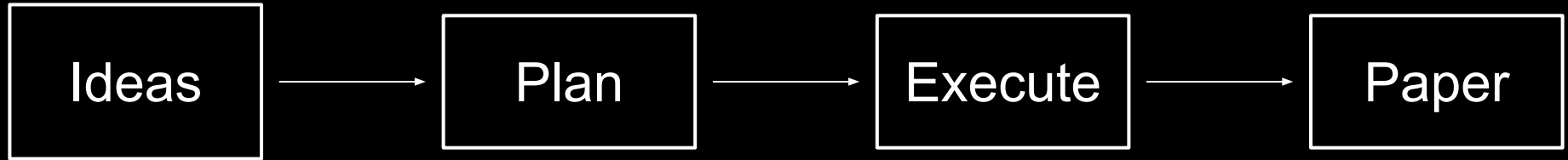
"I'm sure I deserve
a lot of

HOW DO WE WANT TO WORK WITH AI?

Human-AI Collaboration

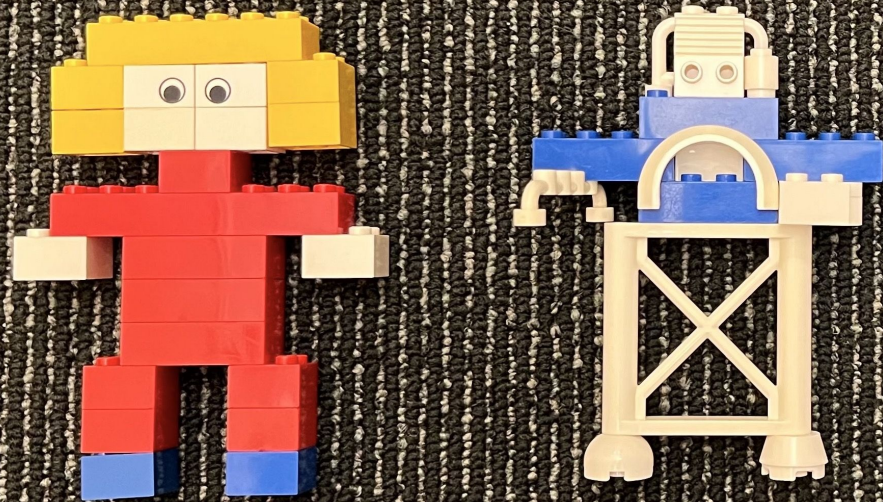


WHICH IS YOUR FAV?



Currently open for participants: <https://cd3.ipmu.jp/hac/>
Perk (TBC): ChatGPT Plus subscription

“Human-AI Collaboration”



Organizing and advising team:

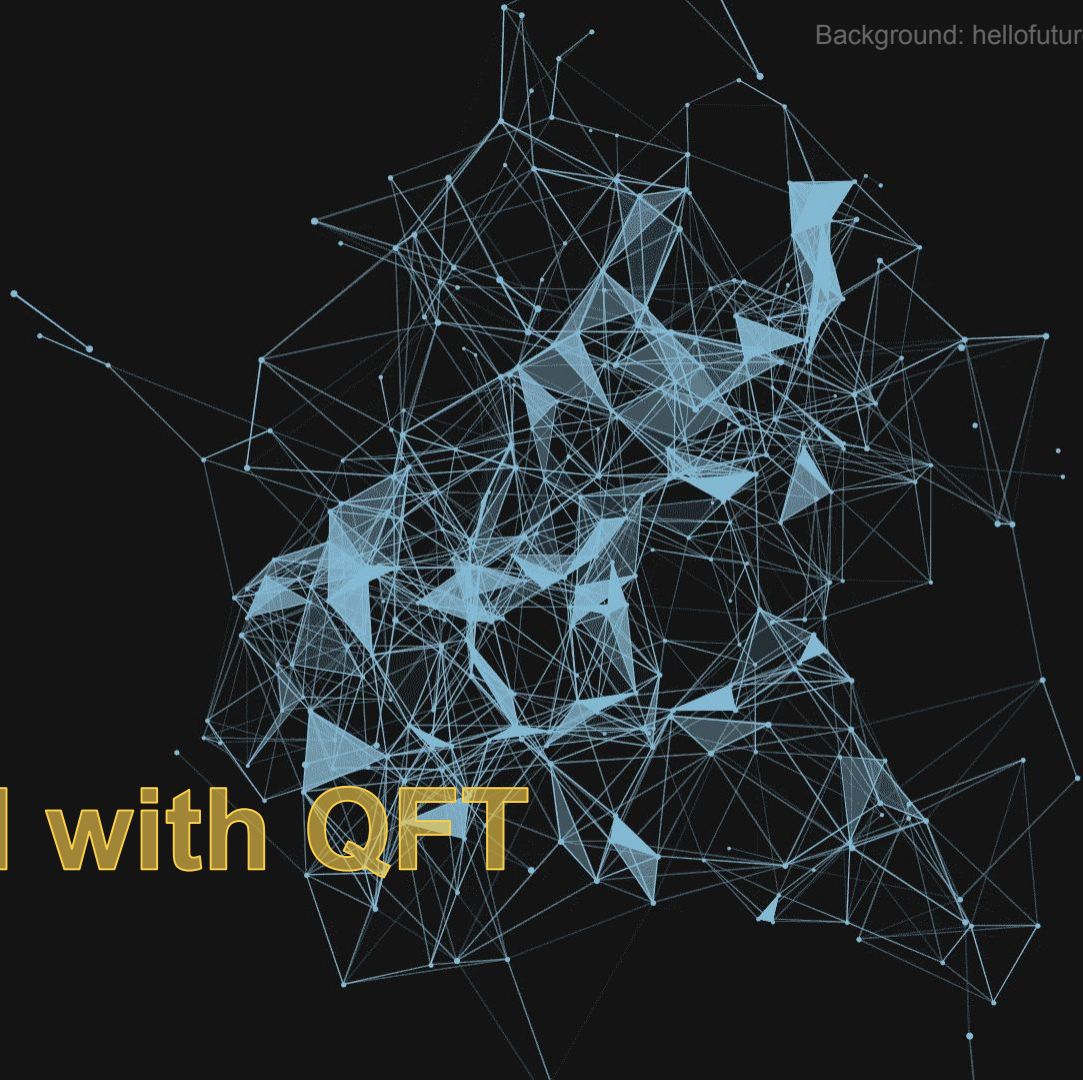
Ben Horowitz, Leander Thiele, Hiromi Yokoyama

Sherry Song, Kateryna Vovk, César Jesús-Valls

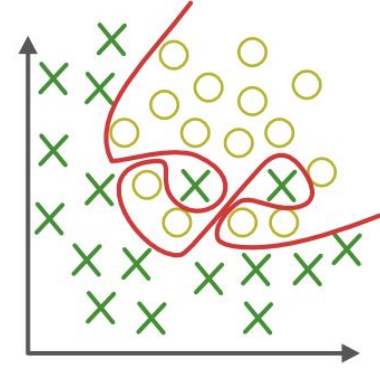
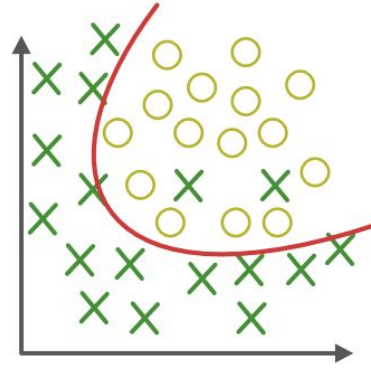
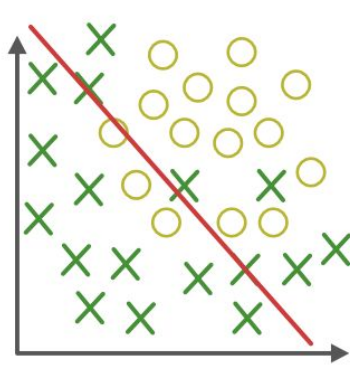
Joaquin Armijo, Sam Passaglia¹, Federico Trucchia²

[1] Lead Machine Learning Engineer, ELYZA Tokyo; Former IPMU PD

[2] DLX Design Lab Co-Director, The University of Tokyo



DECODE AI with QFT

**Under-Fitting****Right-Fitting****Over-Fitting****Training Error**

High

Low

Low

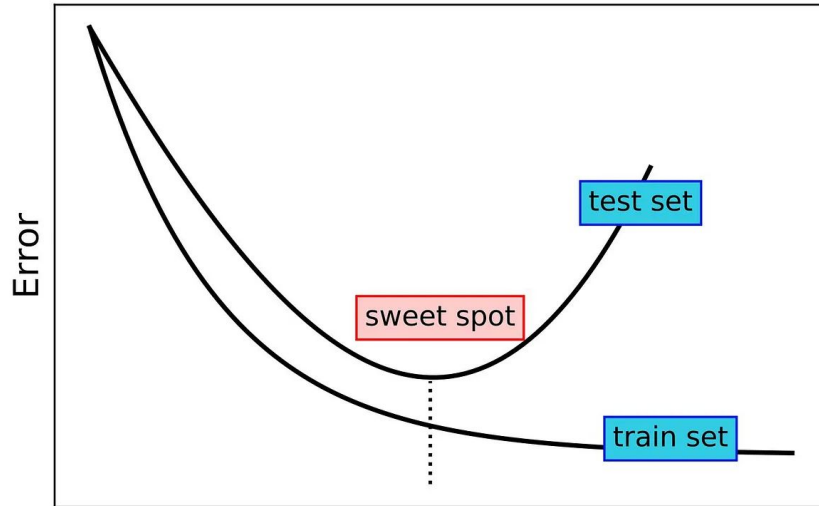
Test Error

High

Low

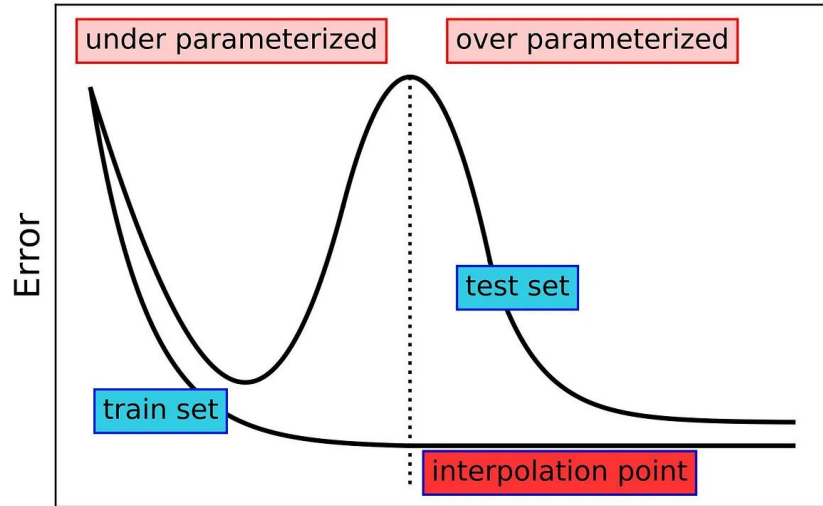
High

Occam's Razor: the simpler the better



Model Complexity

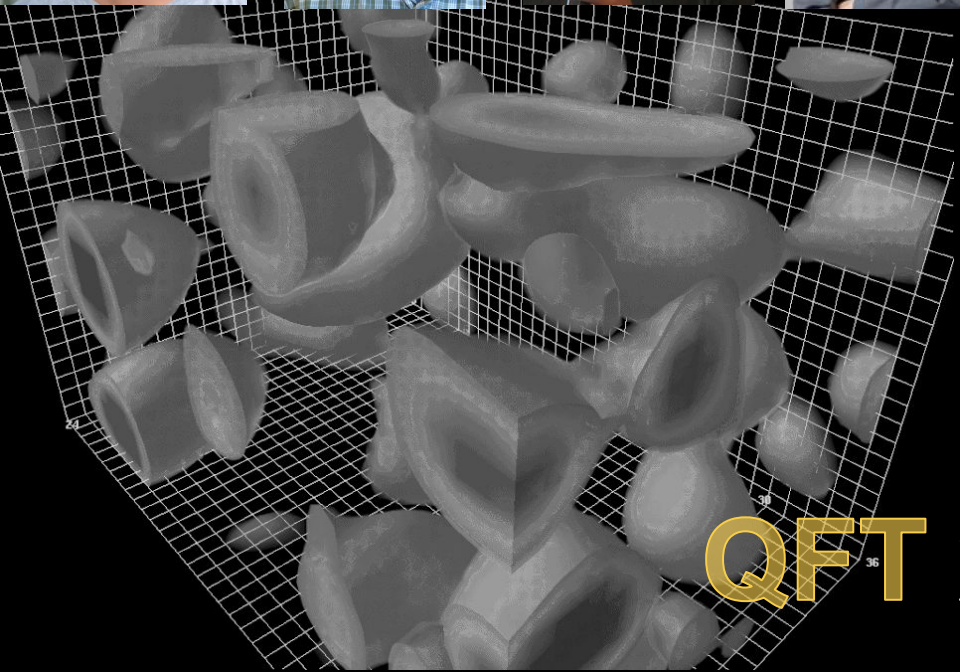
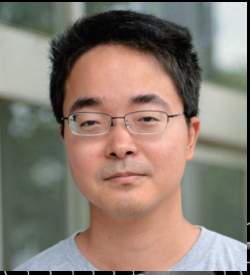
But... ML says otherwise



Model Complexity

E.g. **AlexNet** (2012 ImageNet winner): **60 million** parameters trained on **1.2 million** images

Currently in conversation: Masahito Yamazaki, Simeon Hellerman, Tom Melia, Jun'ichi Yokoyama, Elisa Ferreira, Leander Thiele. Anyone is welcome to discuss!



QFT x ML

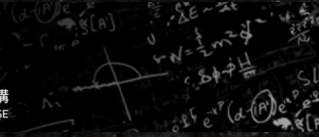
CONCLUSION





KAVLI
IPMU

東京大学 国際高等研究所 カブリ数物連携宇宙研究機構
KAVLI INSTITUTE FOR THE PHYSICS AND MATHEMATICS OF THE UNIVERSE



Phenomenology

Quantum Field Theory

String Theory

Mathematics

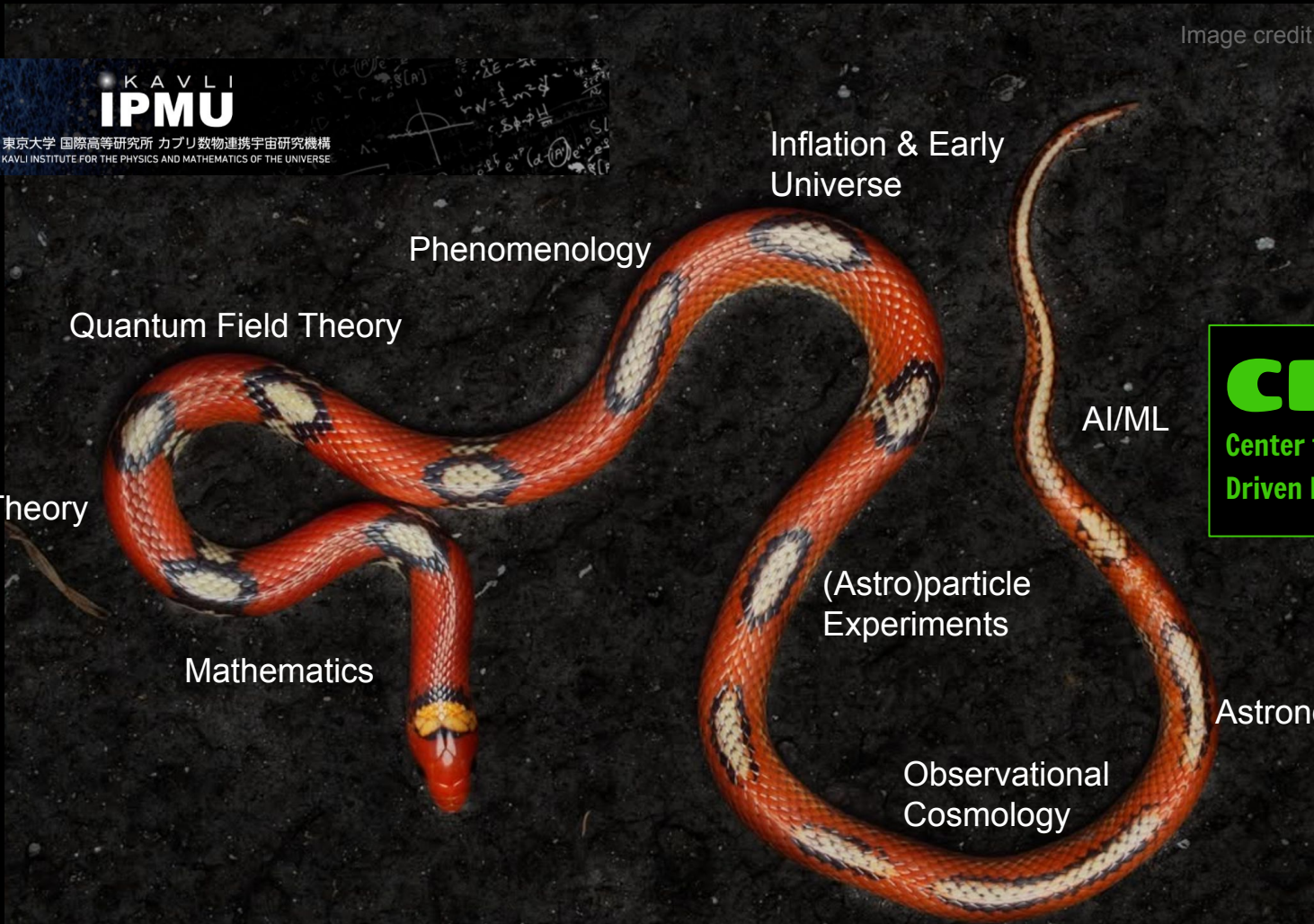
AI/ML

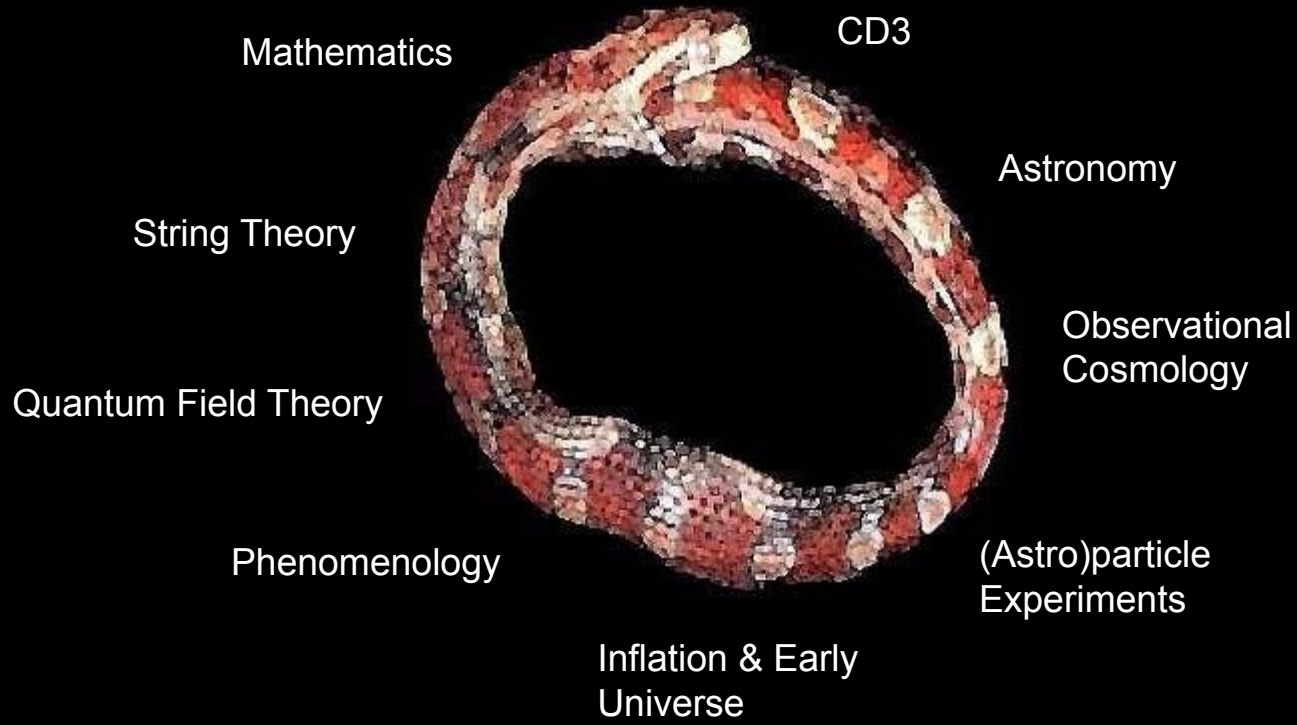
CD3
Center for Data-Driven Discovery

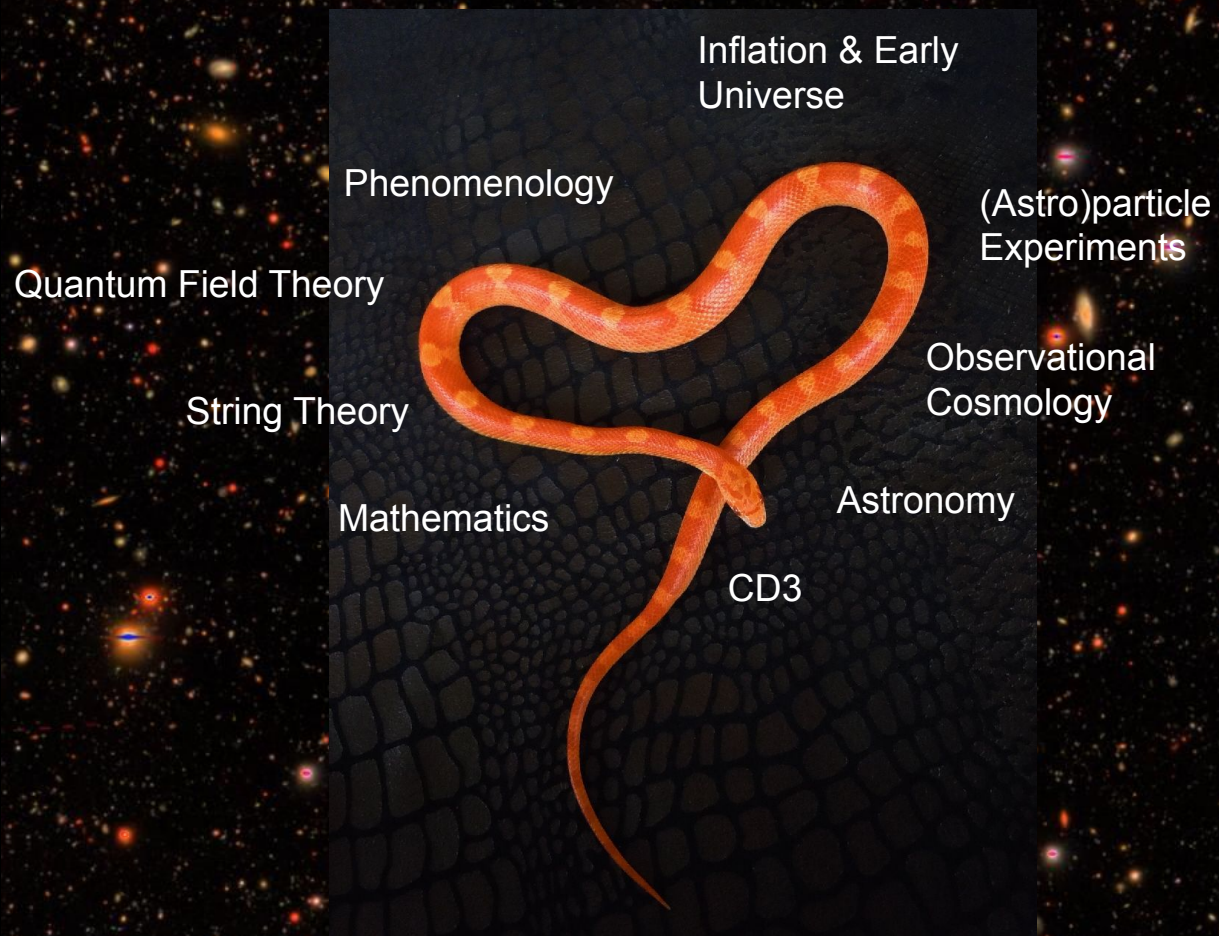
(Astro)particle Experiments

Astronomy

Observational Cosmology







Thank you Hitoshi, for pouring your love and life to build IPMU and the scientific community!



Teaching at the Fundamentals Summer School, Split, Croatia, Sept 2024