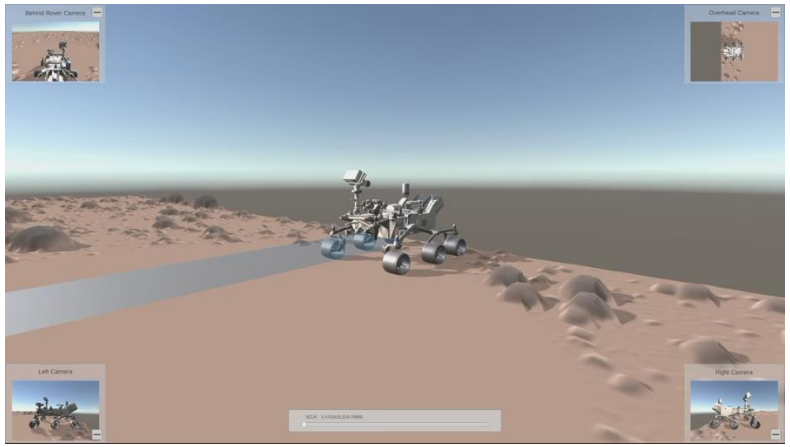
Caltech

Adaptive Robotics Technologies

Intelligence

Yisong Yue

AI for Decision Making



www.yisongyue.com

@yisongyue

Foundations Models for Science

AI Paradigms

More Efficient Task-Specific Decoding

Supervised Learning

Labeled Data

AI Model

Predictions

Pre-Training + Fine-Tuning

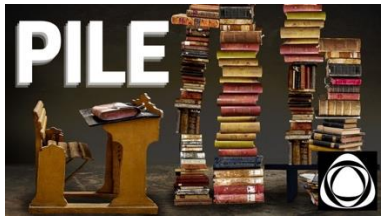
Large Background Dataset

Self-Supervised

Embedding

Fine-Tune on Supervised Data

Foundation Models



Internet-Scale Dataset

Pre-Training + Alignment

Foundation Model

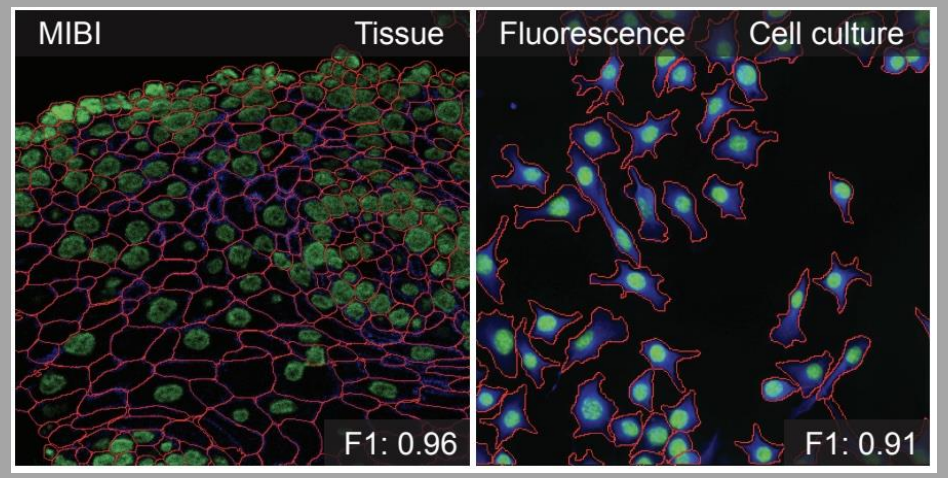
Instructions

Predictions

AI Paradigms

More Efficient Task-Specific Decoding

Can now build AI models for many tasks!



Science

https://cellsam.deepcell.org/

Autonomy

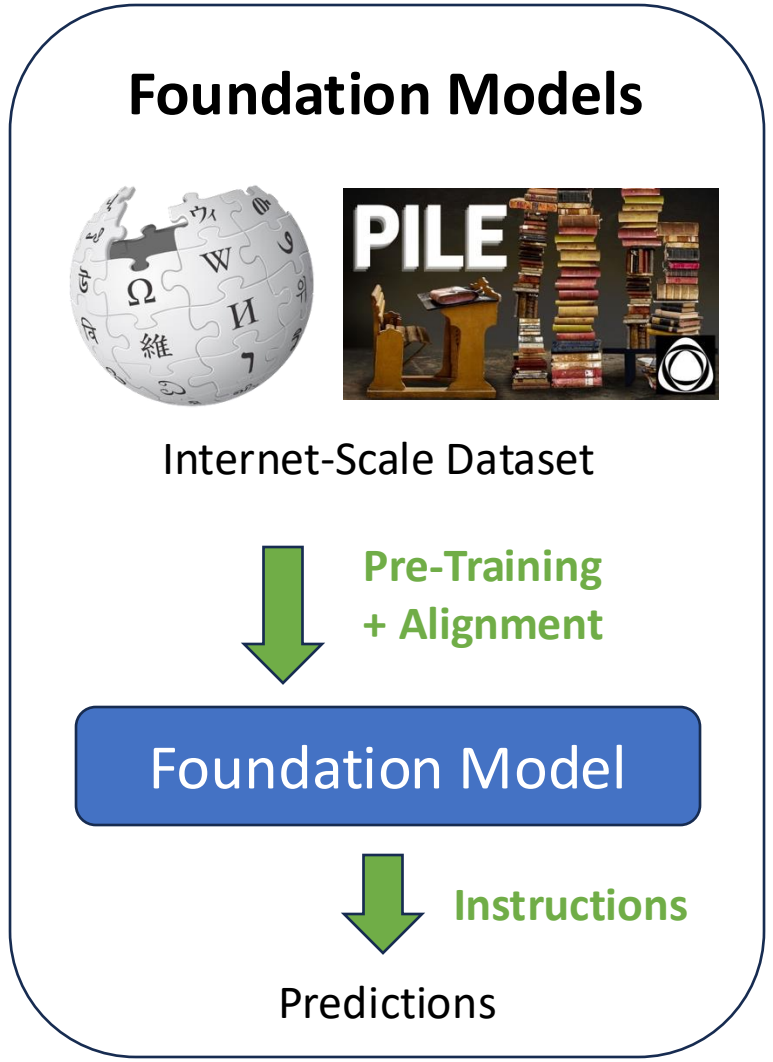
https://www.zuken.com/us/blog/how-are-satellites-bringing-low-latency-internet-to-autonomous-vehicles/

Knowledge Work

(Github Copilot)

What are Foundation Models?

- Reusable for many tasks
- Efficient adaptation
  - Sometimes with just language instructions
  - Sometimes training small decoder
  - Sometimes fine-tuning
- Do not start from scratch



Beyond Language: Vision Foundation Models

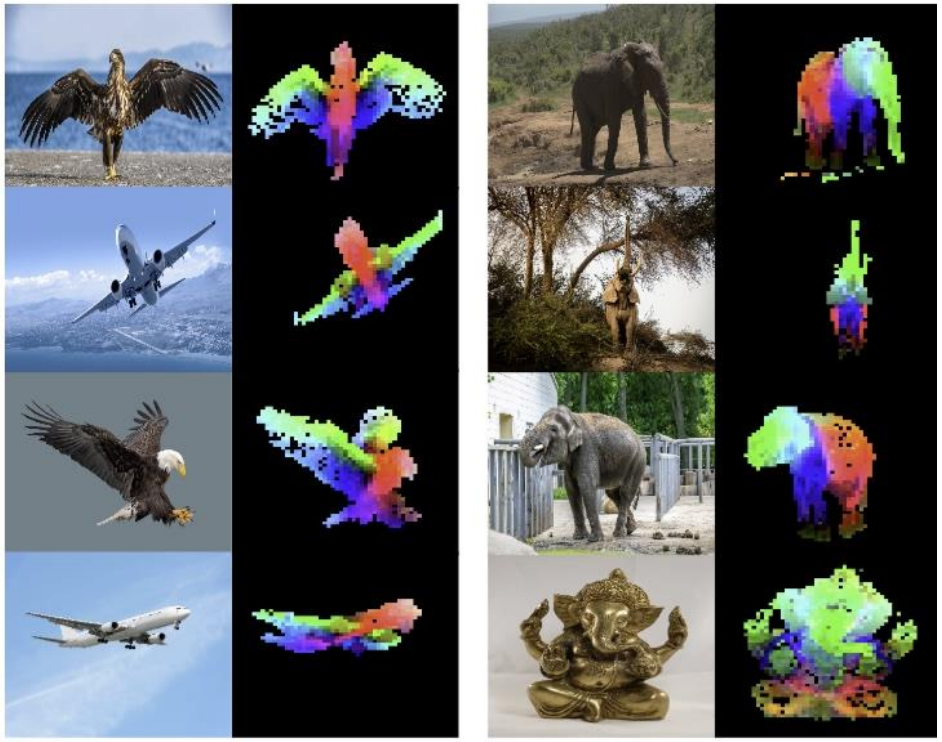
Large Background Corpus
(Careful Data Cleaning)

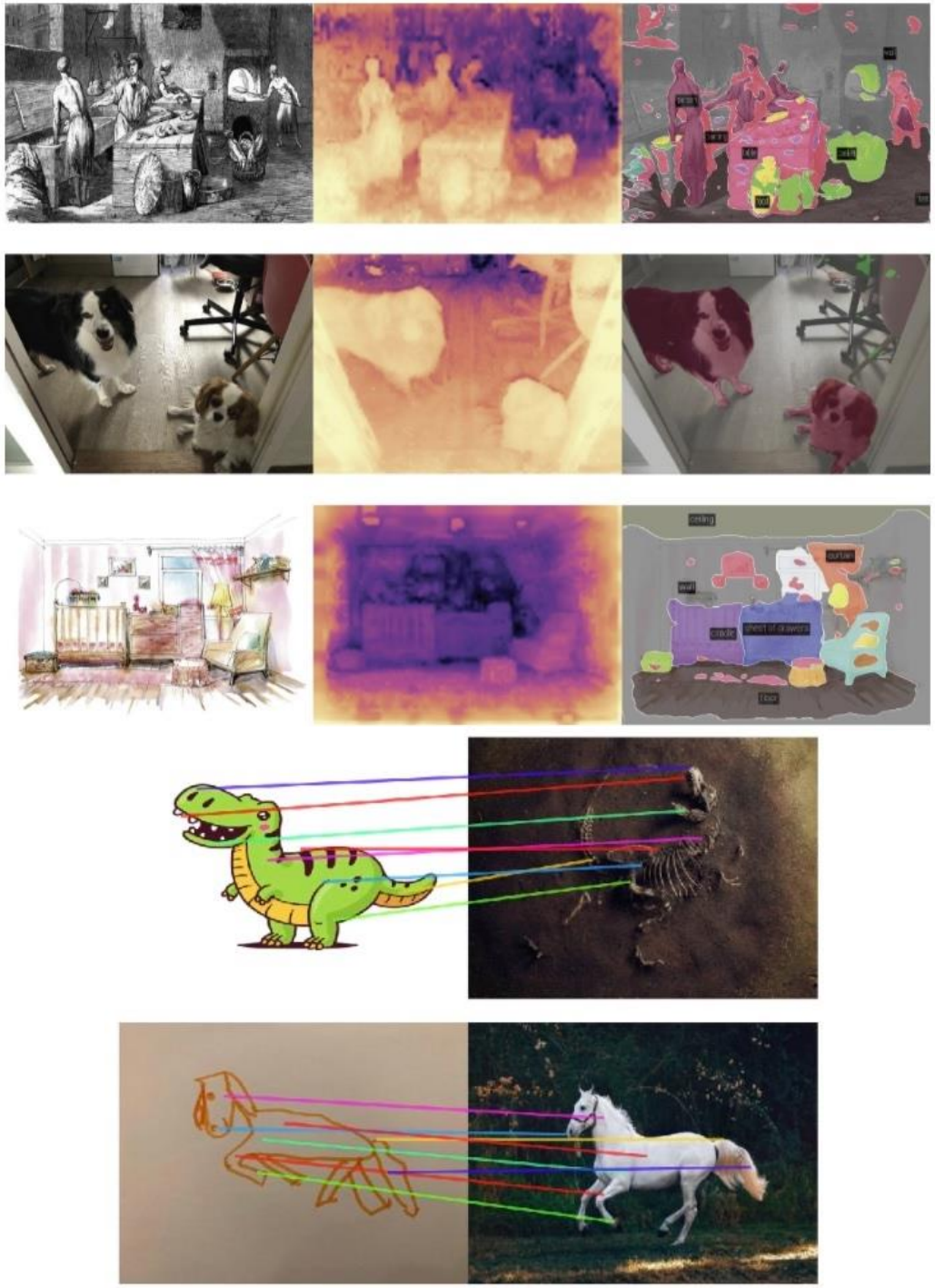
Self-Supervised Learning

Embedding

Many Tasks!

Dino V2 (from Meta AI)

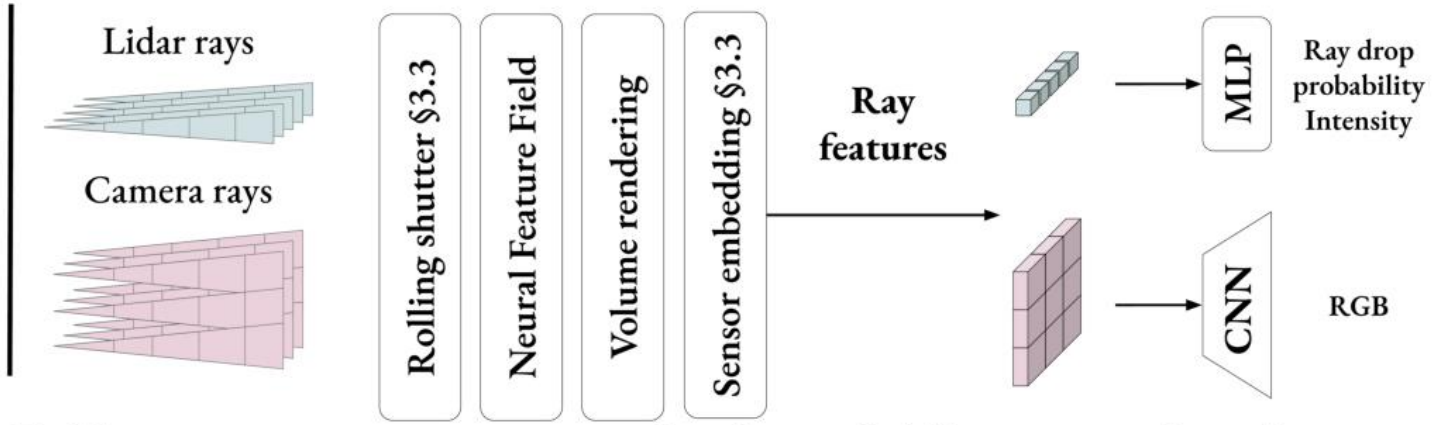


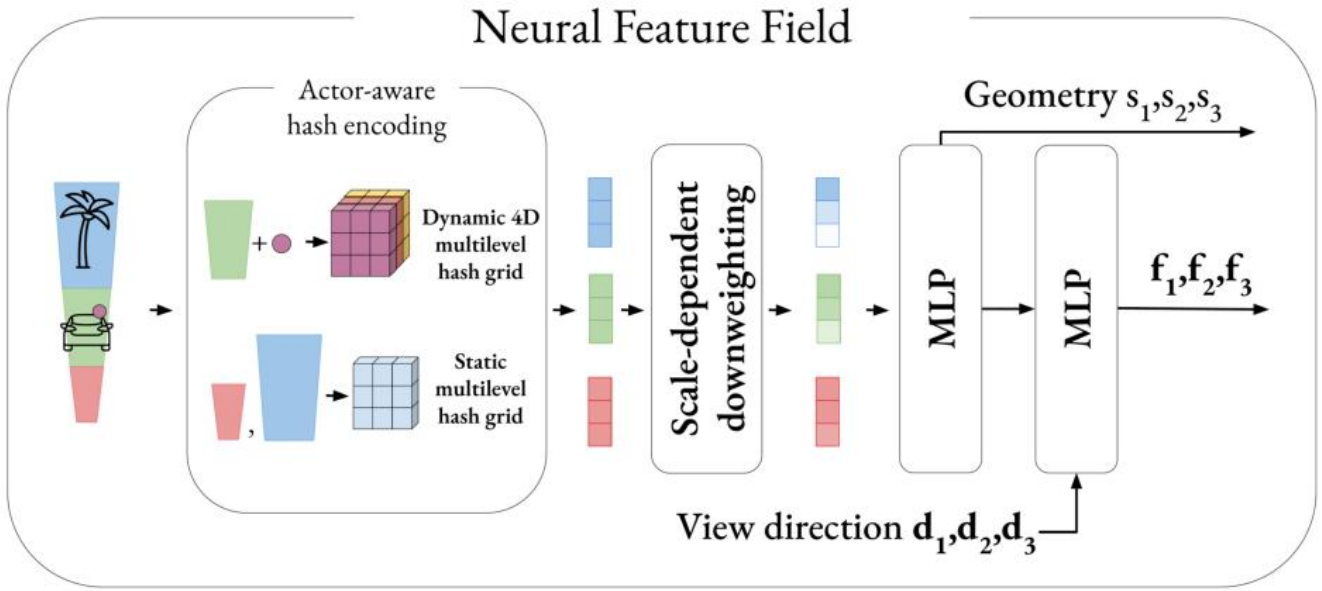


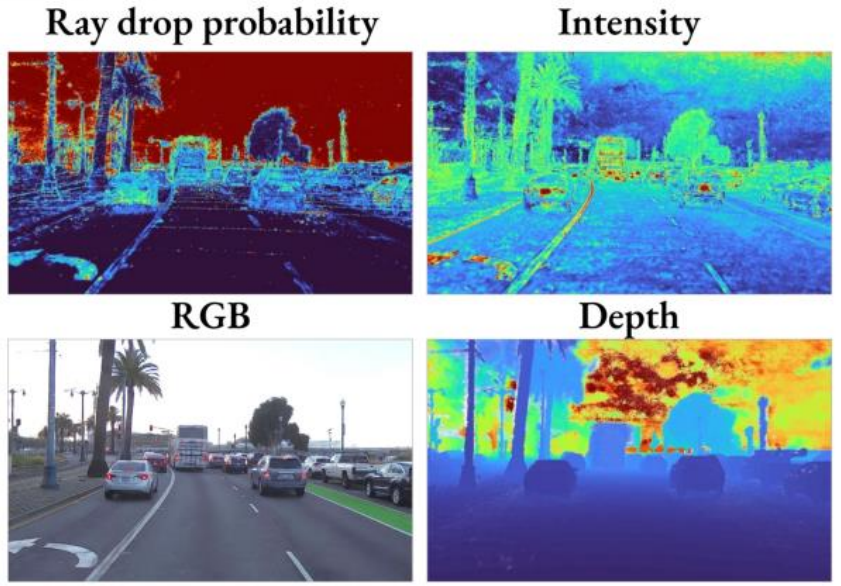
https://arxiv.org/abs/2304.07193

World Models for Autonomy

https://research.zenseact.com/publications/neurad/







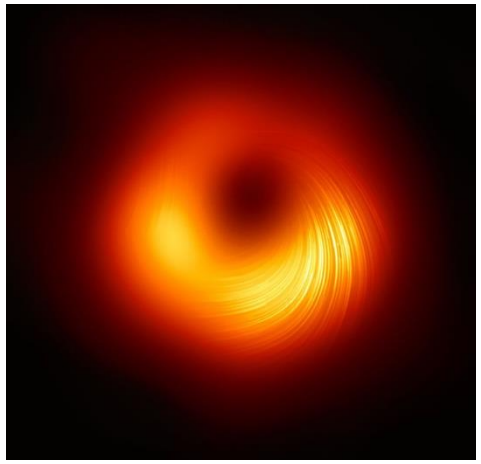
World Models for Autonomy

https://research.zenseact.com/publications/neurad/

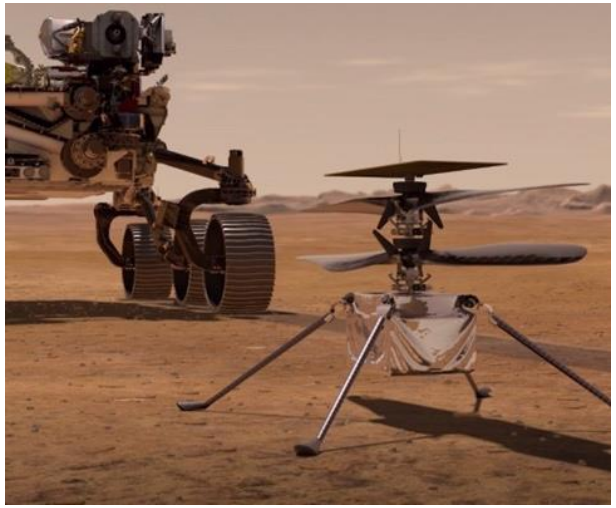
Real Systems have Complex Requirements

"I want to use deep learning to optimize the design, manufacturing and operation of our aircrafts. But I need some guarantees."

– an Aerospace Director while visiting Caltech



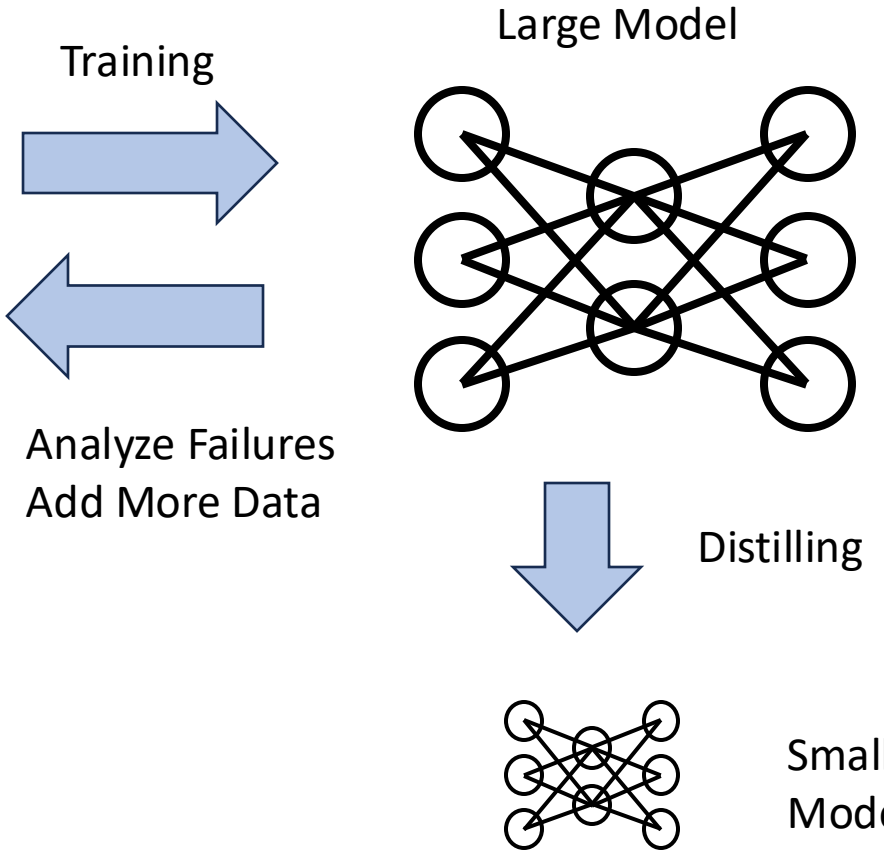
Valid Inferences



Situational Awareness

Operations & Compliance

The First Key to Scaling AI is to Let it Scale!



Tera AI

Tony Zhang

CEO

Google X AI Lead

Caltech Neuro-AI PhD

Pure-software, cross-platform localization

3 foundation models work in tandem

to achieve zero-shot generalization across terrains and robots – ***on edge***

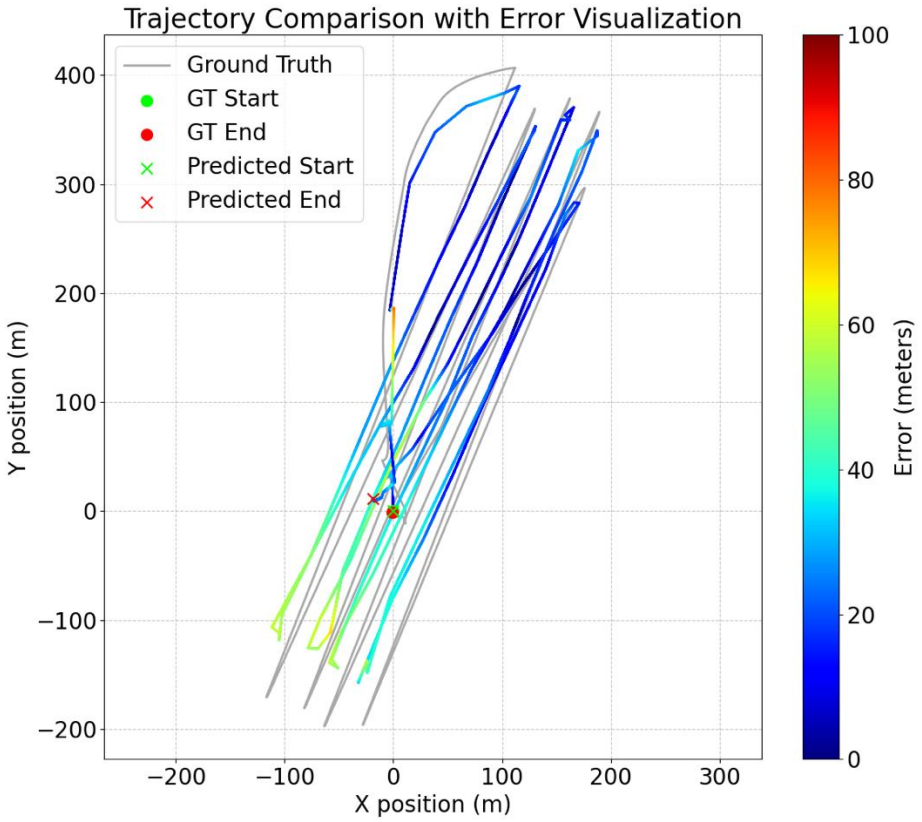
System 1
Deep SLAM

System 2
View-Invariant Map-Matcher

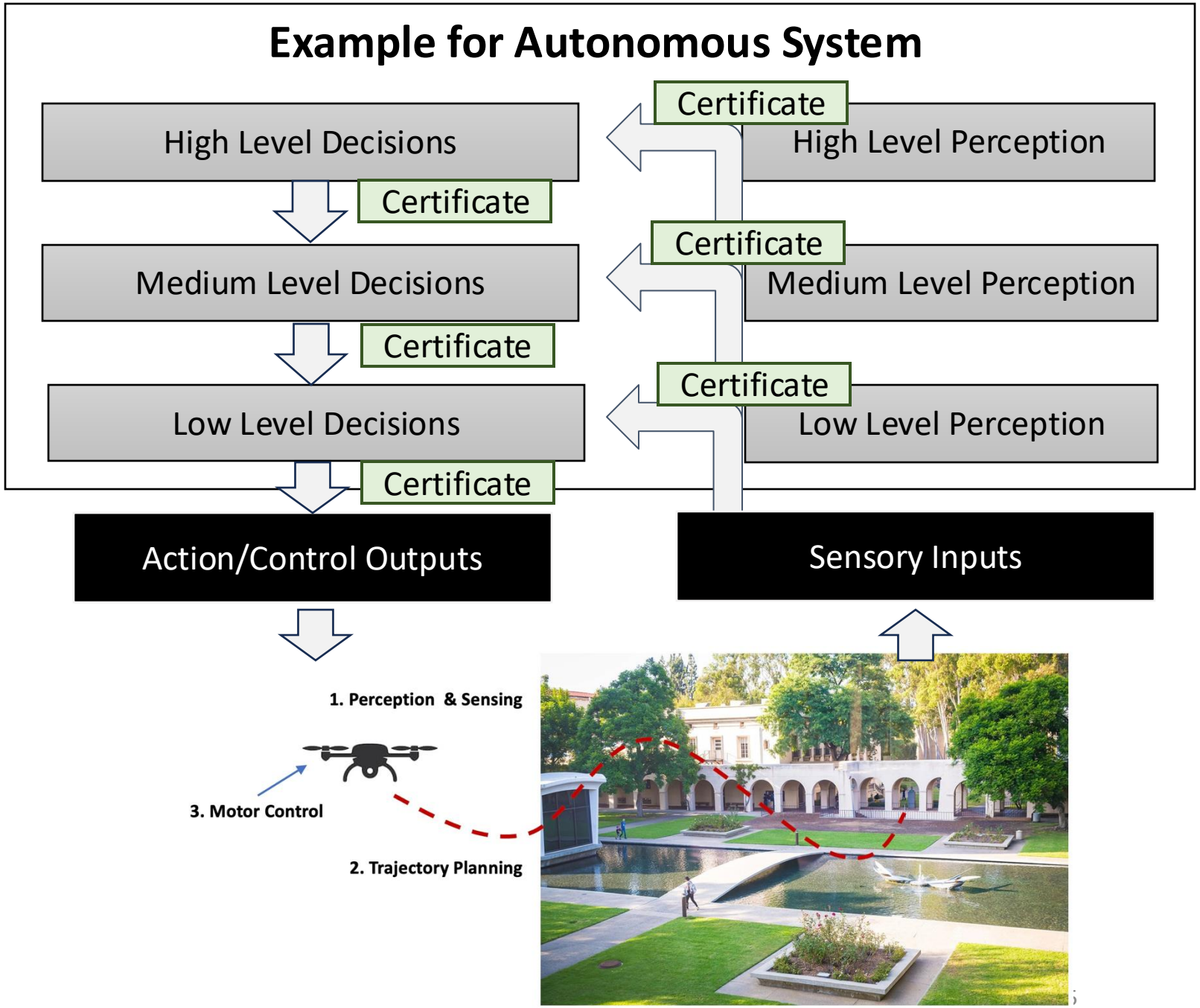
System 3
Global Geoguessr


Validated in the toughest terrains

Localization & Mapping from Monocular Camera



The Second Key is Having Guardrails that are Compatible with Scale



Certificates Define Guardrails!

The Second Key Ge... ar... w...

Autonomous System

Certificate

...tion

...ion

...n

Can use AI to help optimize any aspect!

System-Level Design

↔ End-to-End Learning

→ System Programming

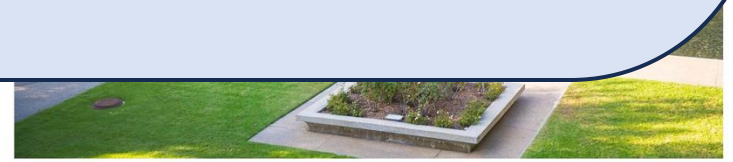
← Library Design

← Learning signal (satisfy certificates)

Component-Level Design

Abstractions & Certificates

→ Deduce what can be certified

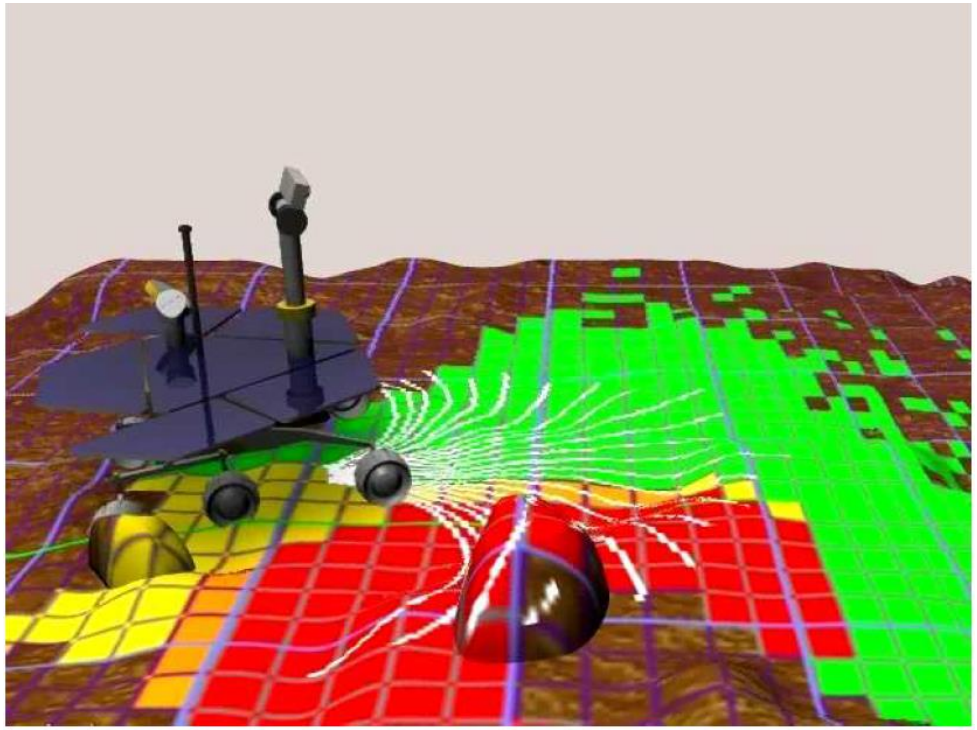


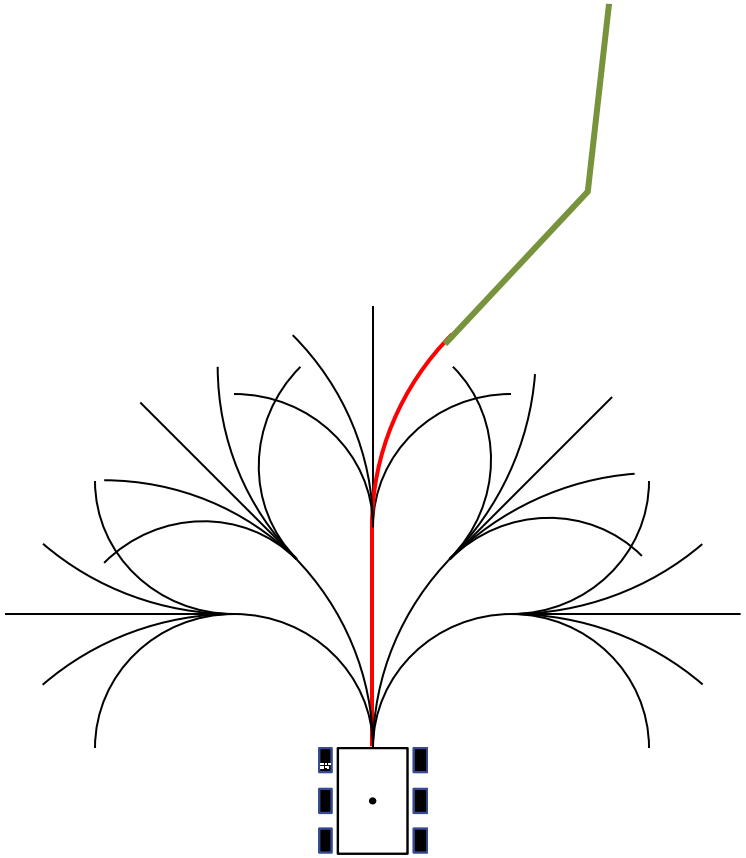
MLNav: Learning-Augmented Rover Navigation


Shreyansh Daftry


Hiro Ono

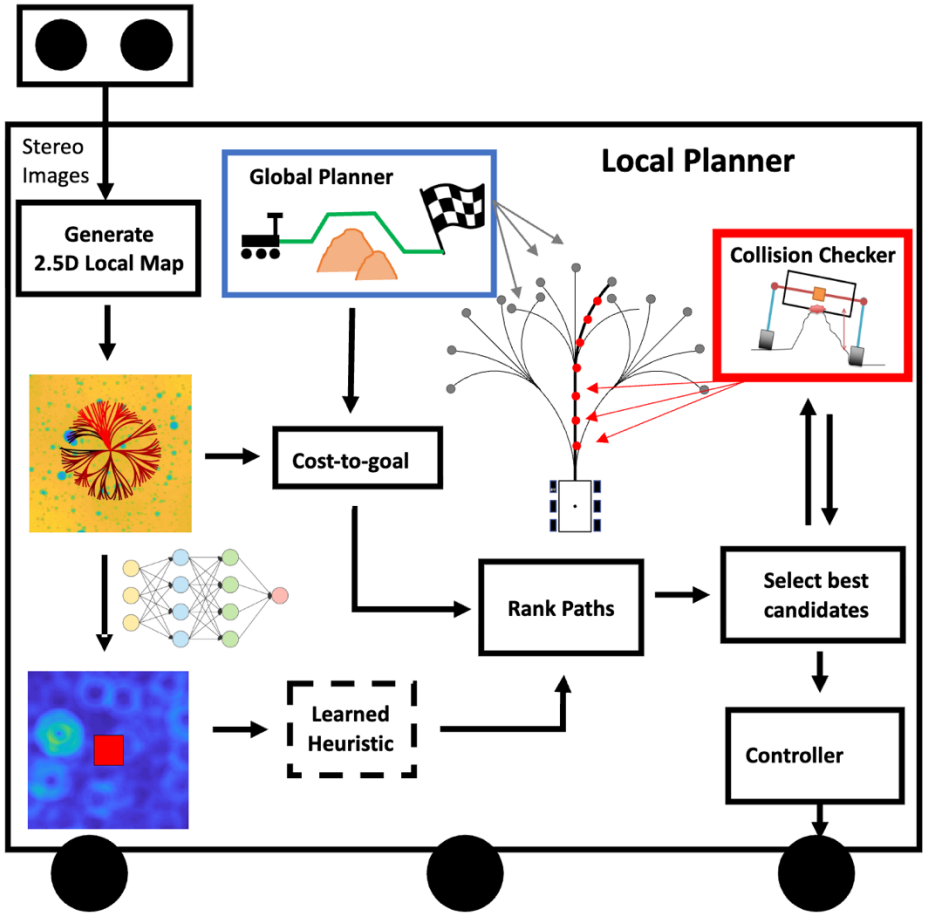

Neil Abcouwer





MLNav: Learning to Navigate on Martian Terrains, Shreyansh Daftry et al., R-AL 2022

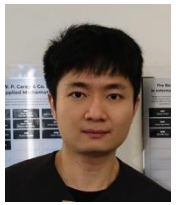
MLNav: Learning-Augmented Rover Navigation



Most Expensive Module
Accelerate with AI

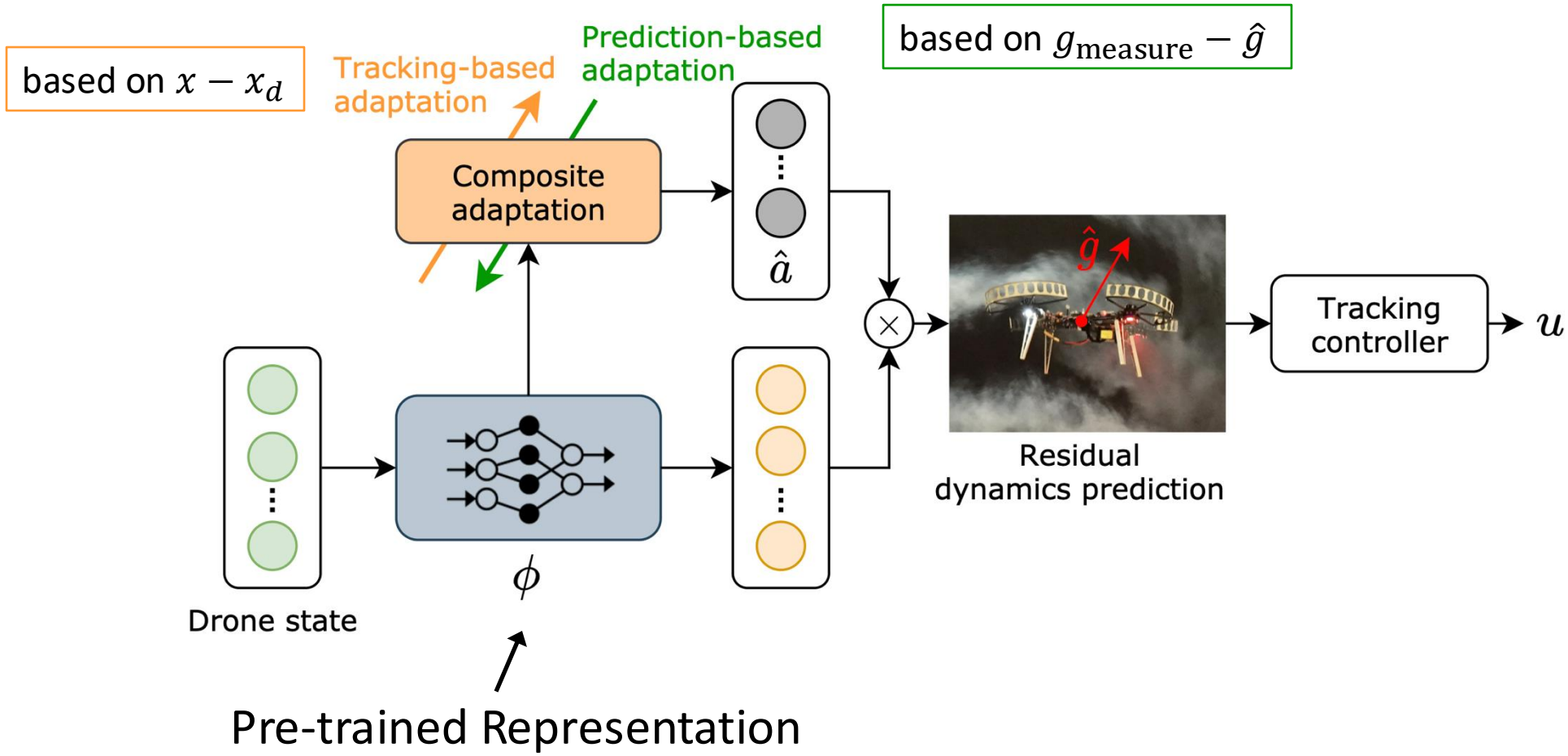

Perseverance Sol 122
Sol = Martian day
Image: NASA/JPL-Caltech

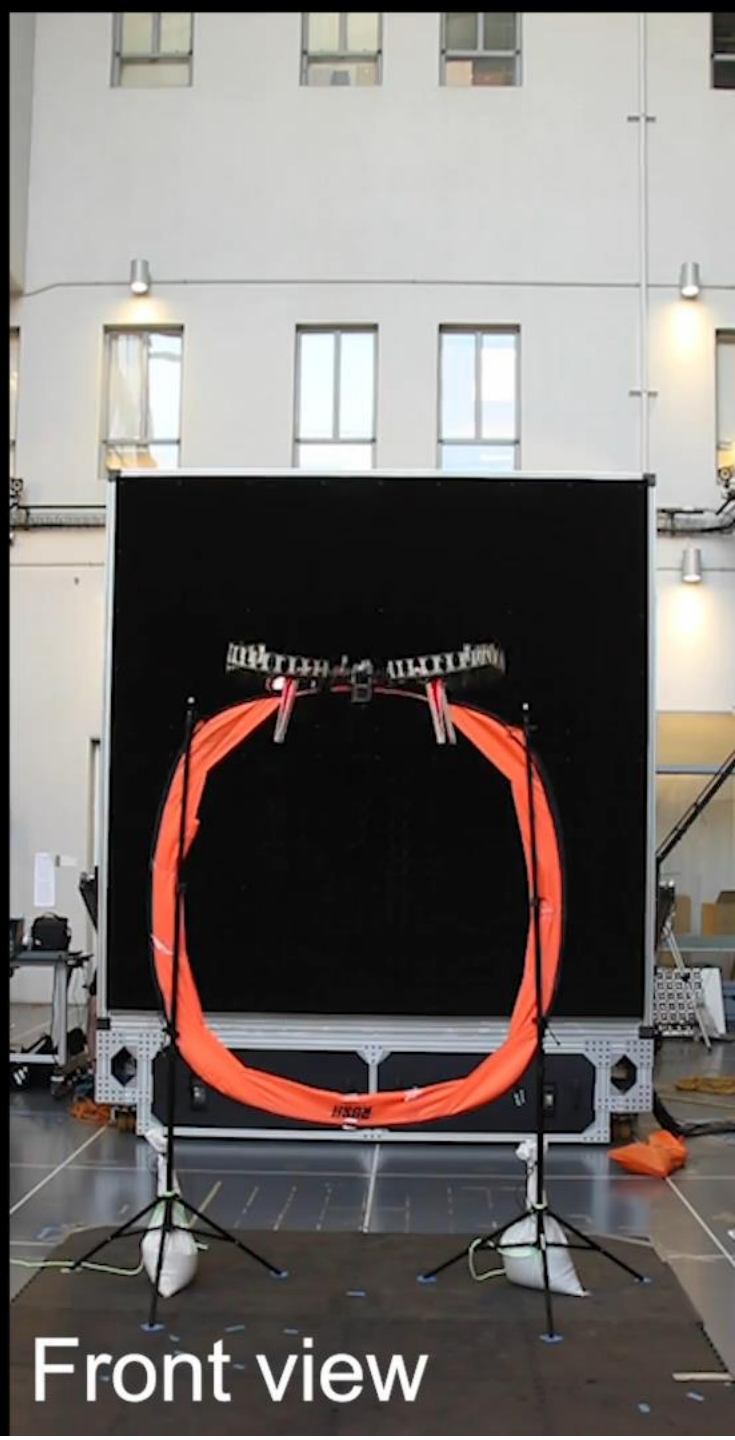
Stable and Robust Adaptive Control


Guanya Shi


Michael O'Connell

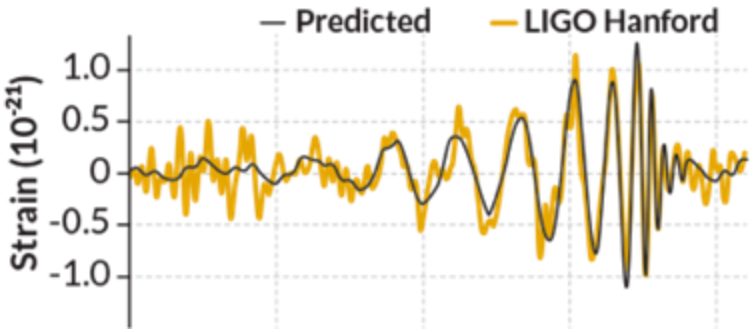

Soon-Jo Chung





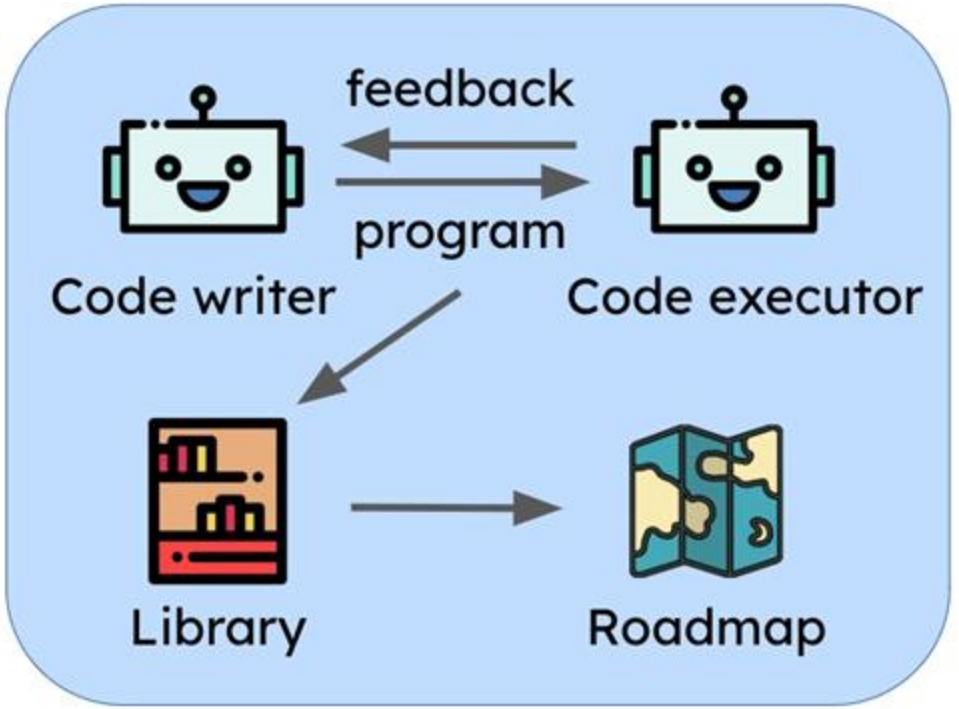
Our method can also generalize to the challenging time-varying wind condition.

Frontier of AI: Self-Improving Agents



Raw Data

Concepts



Learn to use new tools
(complex software, robotic laboratories)

Automated hypothesis search & verification

Hypotheses

$$h = \frac{2G}{c^4}\frac{1}{r}\frac{\partial^2 Q}{\partial t^2}$$

Example: Hypothesis Discovery

Symbolic Regression Example from Miles Cranmer

c. PySR. Miles Cranmer

Ongoing: Discovering New Empirical Laws

ML Problem: discovering non-linear data transformations

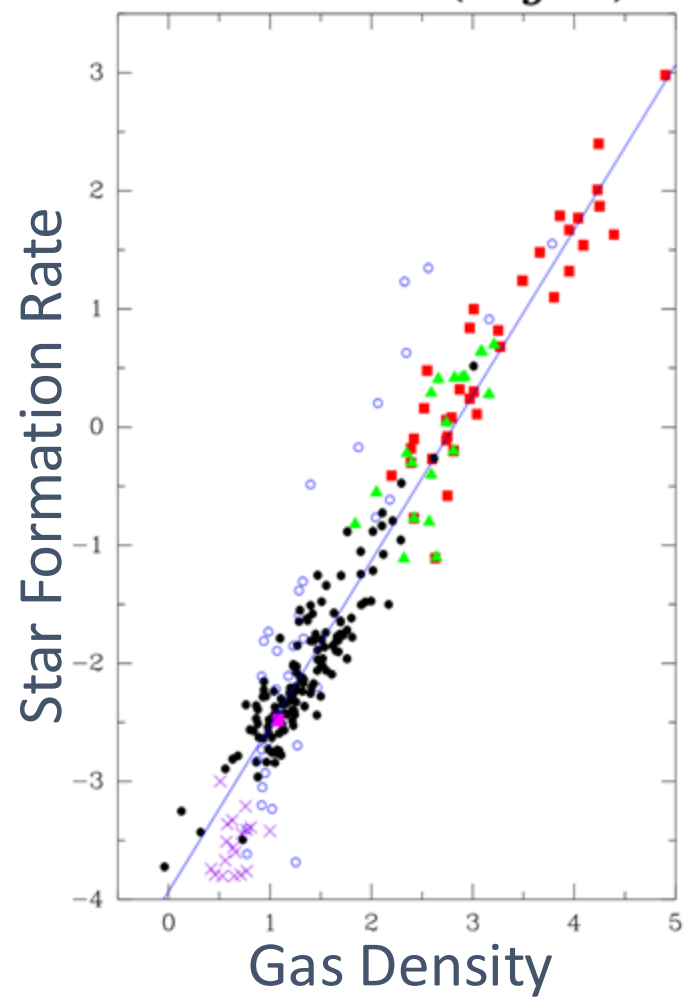
Discovered Law

$$\left(\frac{0.13888639}{0.19667038 + \left(\log\_gas^{\frac{0.0086438805}{\text{Redshift\_NED}}}\right)}\right) - \left(\frac{0.49182796^{\log\_gas}}{0.21173178}\right)$$

redshift

Existing Law

$$\Sigma_{SFR} \propto (\Sigma_{gas})^n.$$



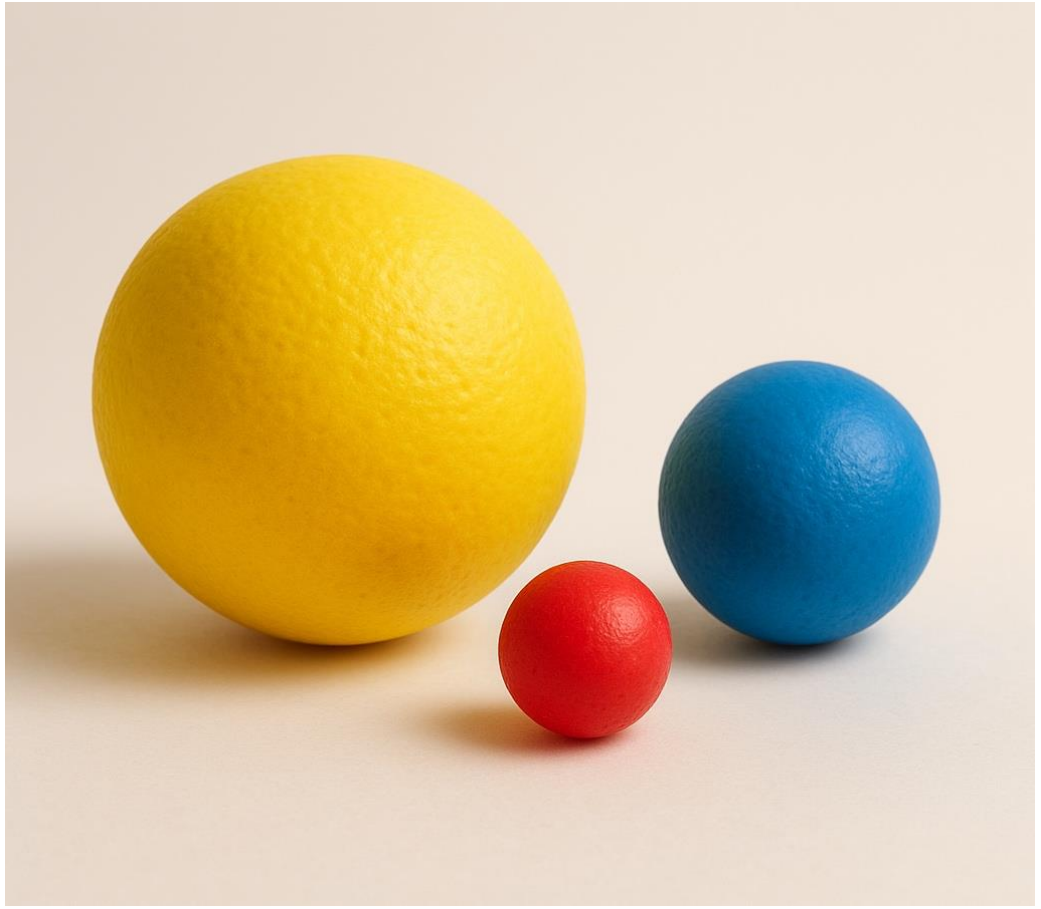
Cosmology Problem: *How does the distribution of matter in galaxies change as they age?*

Frontier of AI: Multi-Modal Reasoning

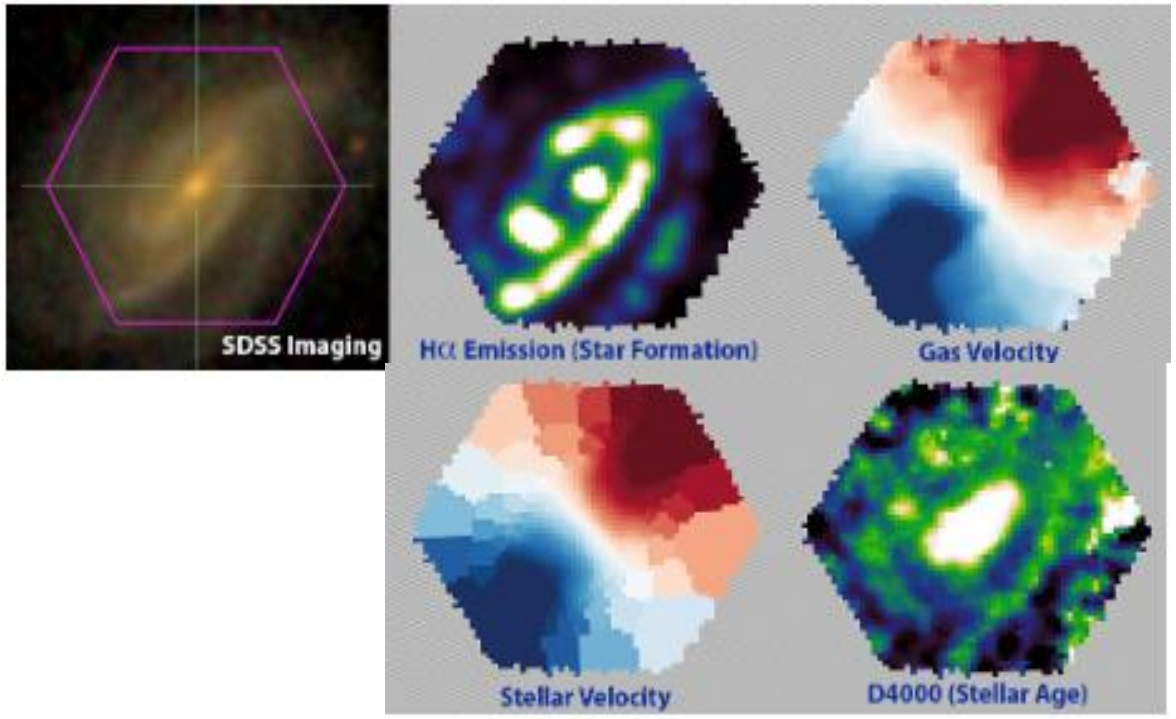
Generate an image of a yellow ball, a blue ball, and a red ball.

The yellow in the front and to the left of a blue ball, from the perspective of the camera viewpoint. The red ball should between the other two.

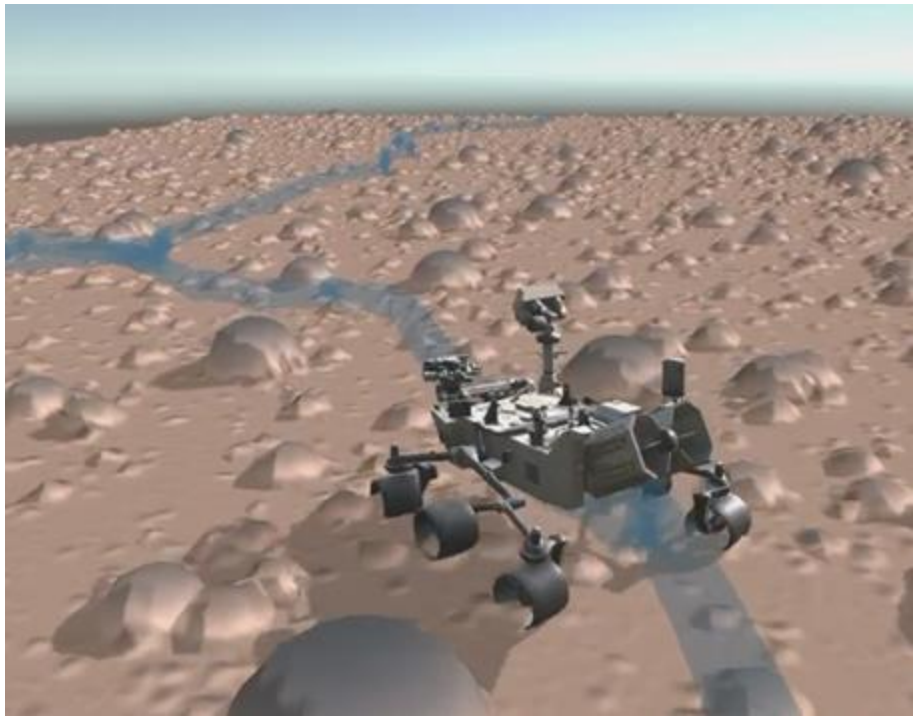
The yellow ball should be twice the size of the blue ball by volume. The red ball should half the size of the blue ball by volume.



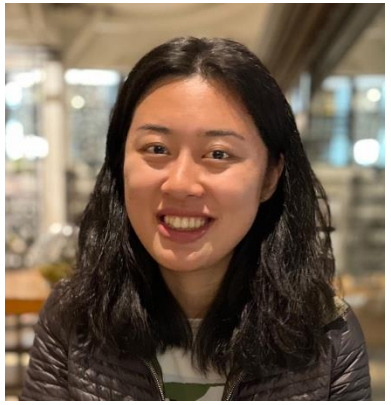
Spatial & Visual Reasoning is Important!



Scientific Analysis & Discovery



Autonomy

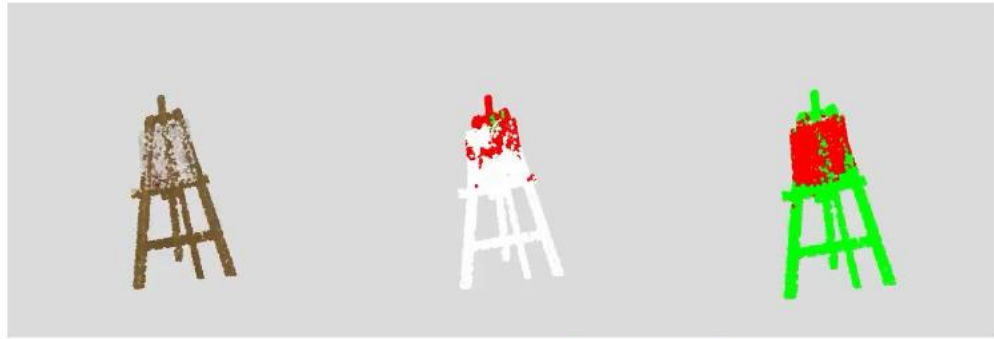


Ziqi Ma

Find3D

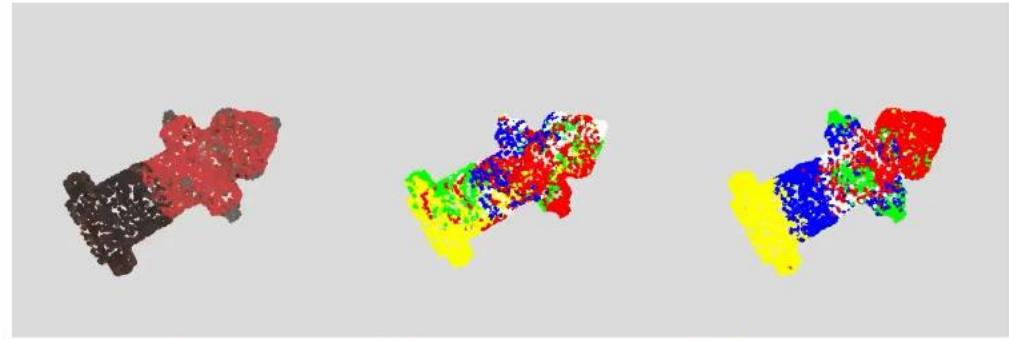
Demo:
https://ziqi-ma.github.io/find3dsite/

Input Existing Method Find3D

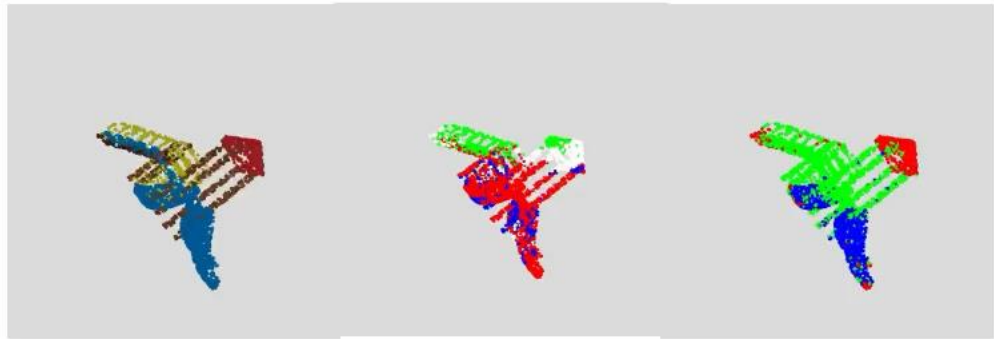


canvas easel stand

Input Existing Method Find3D

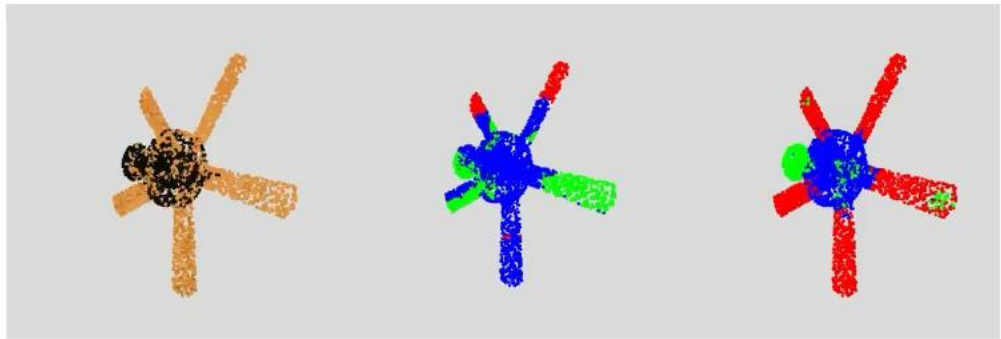


bonnet side cap barrel base

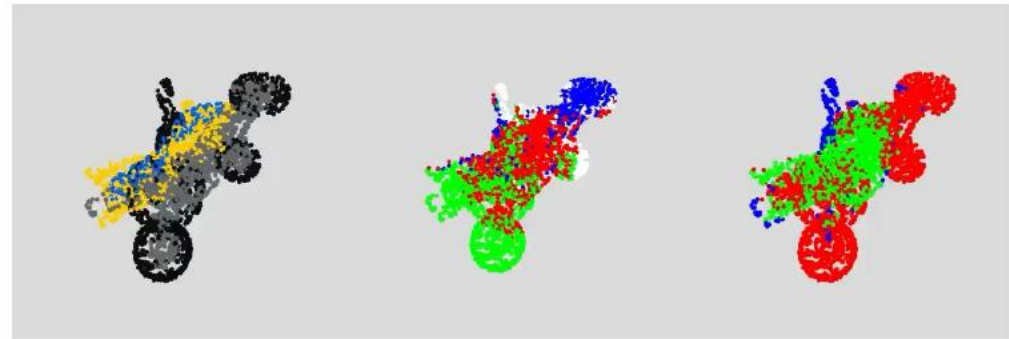


roof staircase slide

sole of a shoe shoe upper



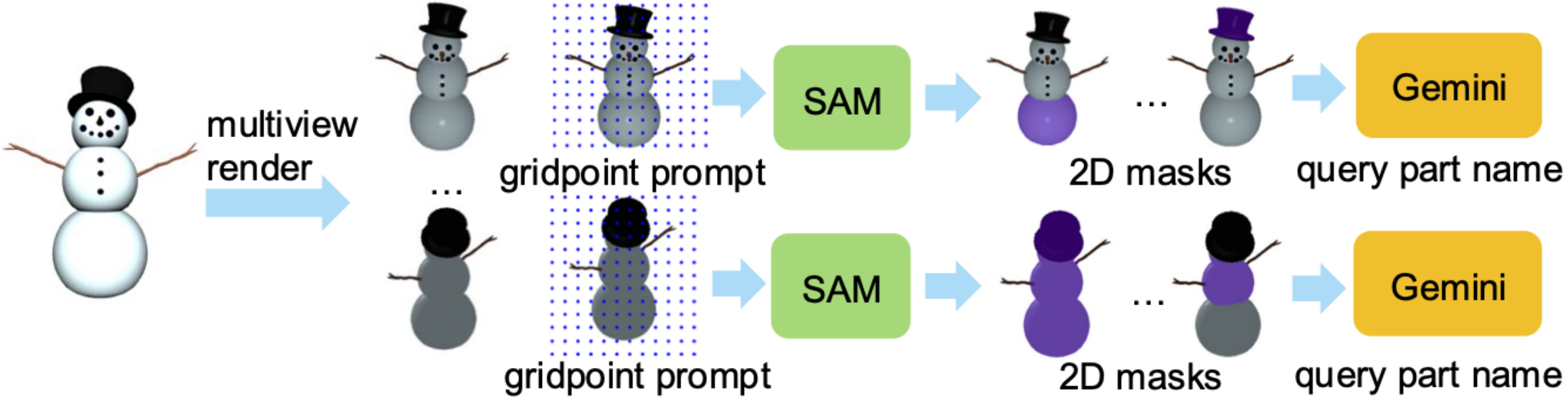
blade motor hub



wheel seat handle

Key Challenge: Data Engine

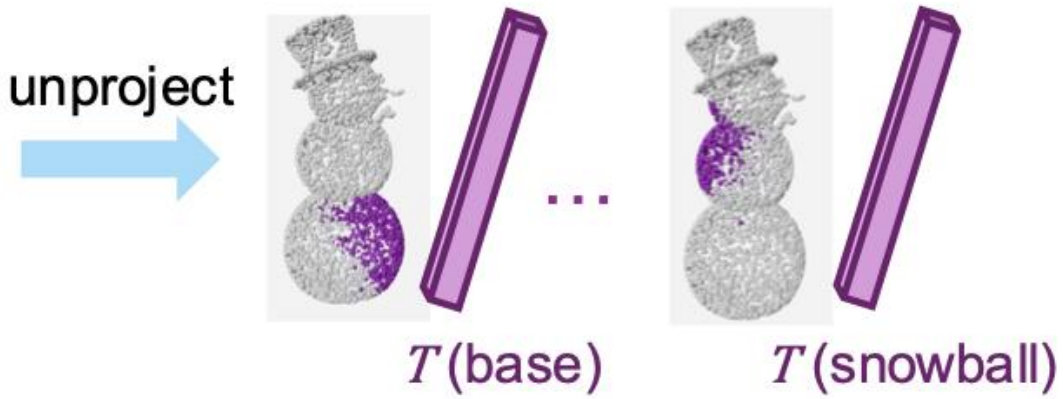
- We have no training data!
- Especially for things in 3D!



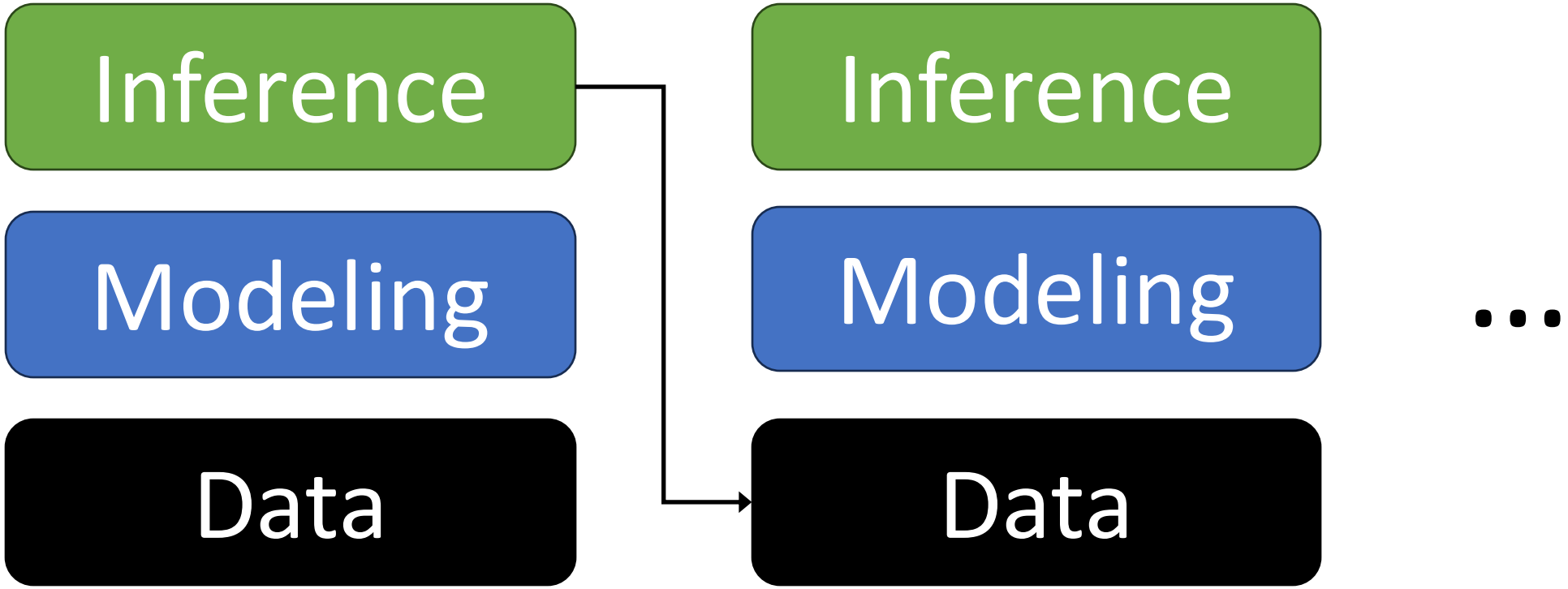
Leverage 2 existing foundation models

- SAM – Segments anything in 2D
- Gemini – Recognize parts in 2D

Create training data in 3D!



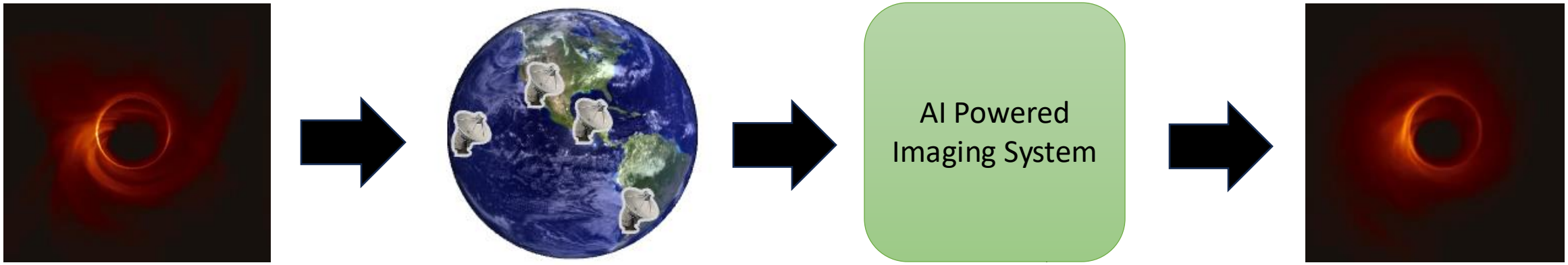
Data Flywheels are Powerful



Frontier of AI: Sensing Systems

Automatically Exploring Massive Hypothesis Spaces

Example: Black Hole Video Reconstruction
(w/ Katie Bouman)

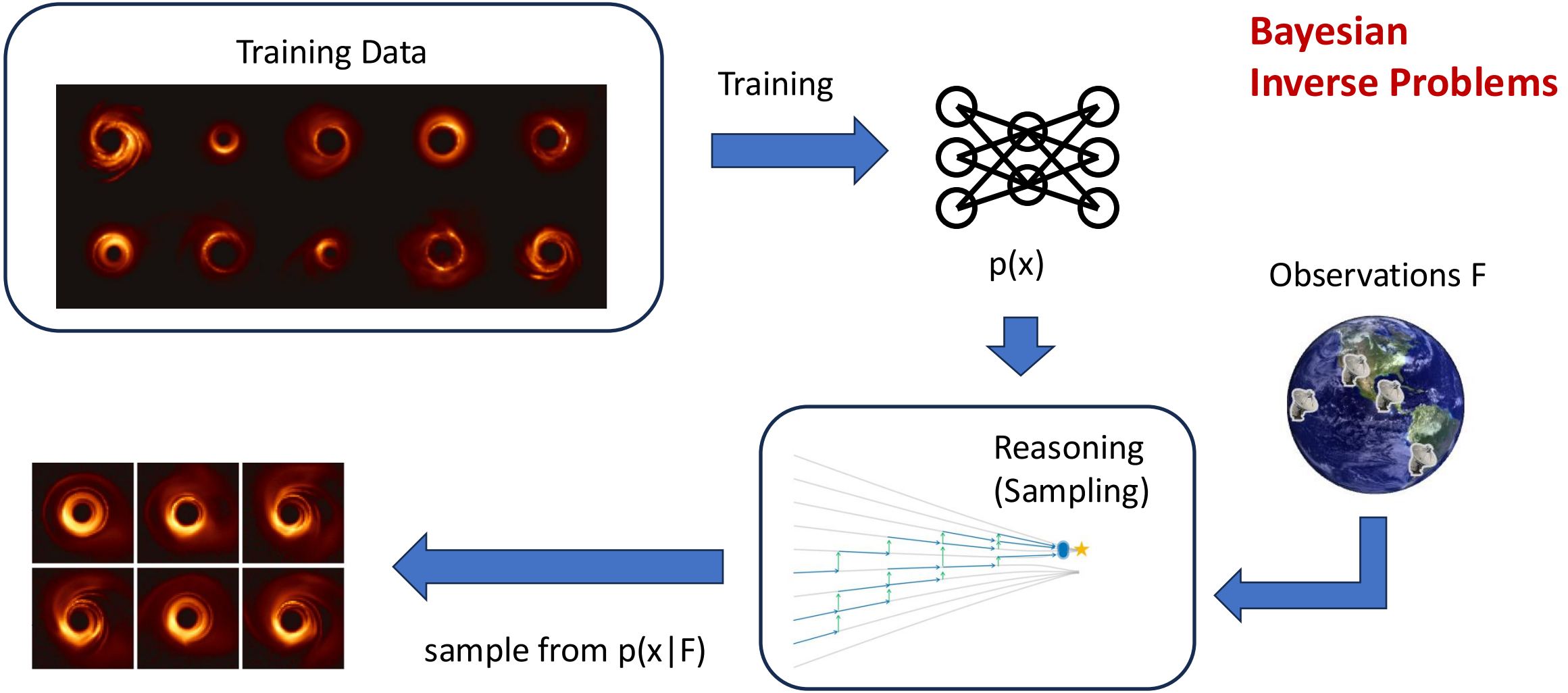


Ground Truth

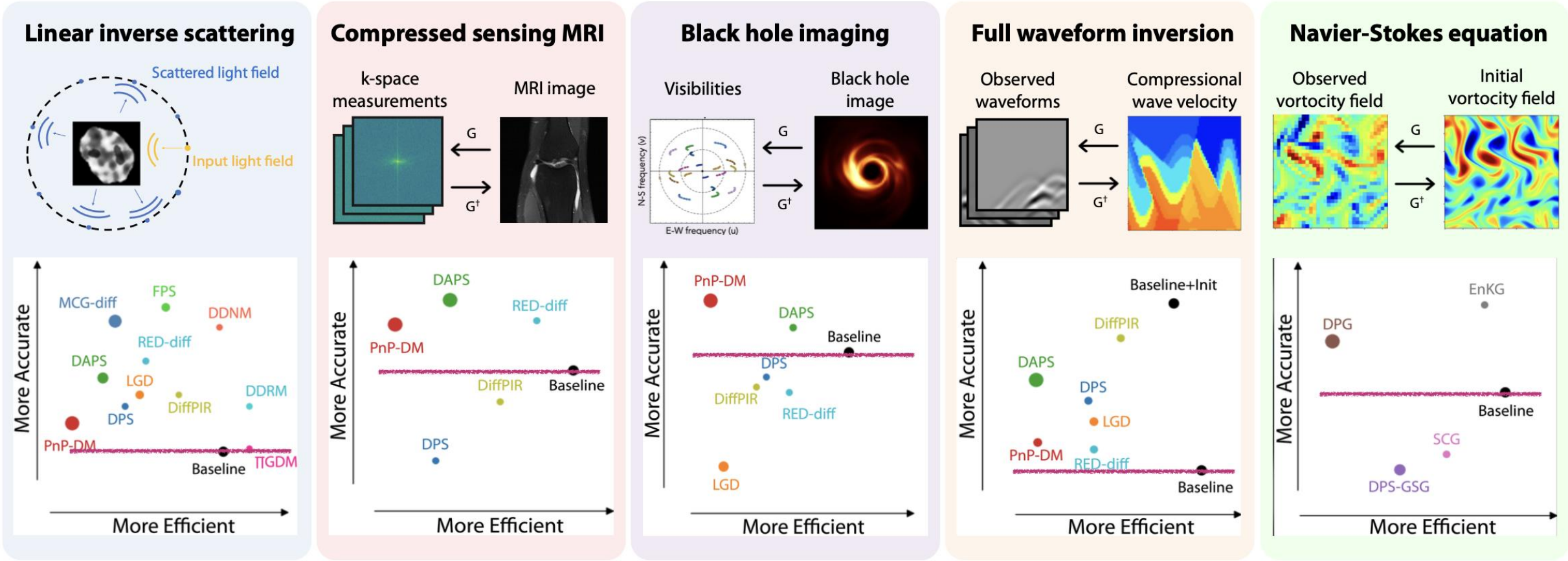
Observations

Reconstruction

Foundation Models for Science

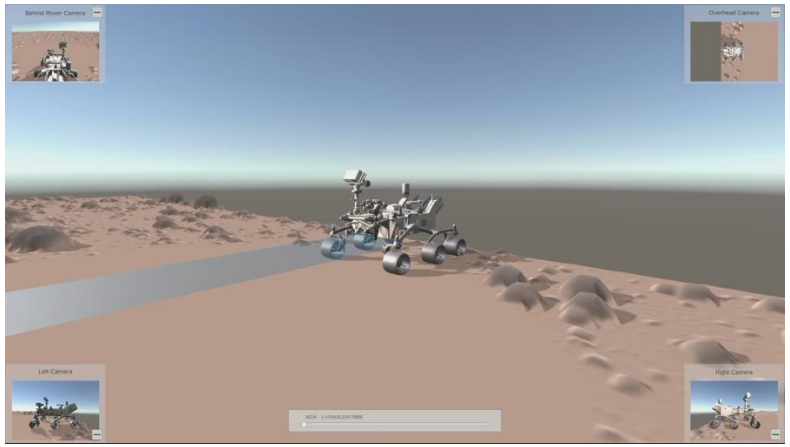


InverseBench



https://devzhk.github.io/InverseBench/

AI for Decision Making



www.yisongyue.com

@yisongyue

Foundations Models for Science