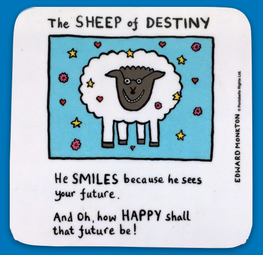
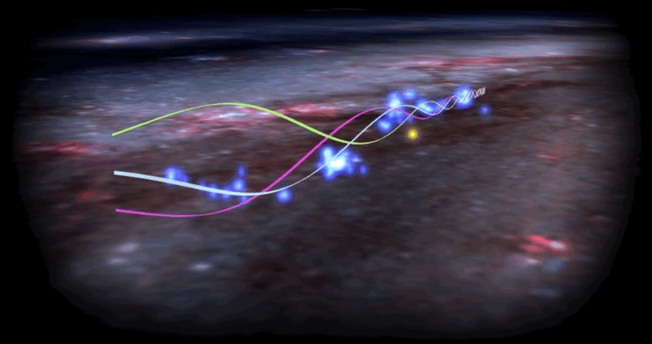
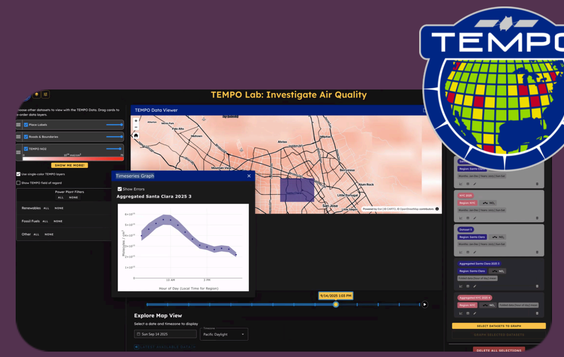
How Mesopotamian Sheep Entrails, Stars forming in the Milky Way, and Climate Change are Related





glue
multidimensional data exploration



PREDICTIONX

MILKYWAY3D.org

LIVE

CosmicDS

Prof. Alyssa A. Goodman, Center for Astrophysics | Harvard & Smithsonian

What could this mean for NVIDIA + PredictionX/MilkyWay3D/LIVE?

CosmicDS

PREDICTIONX

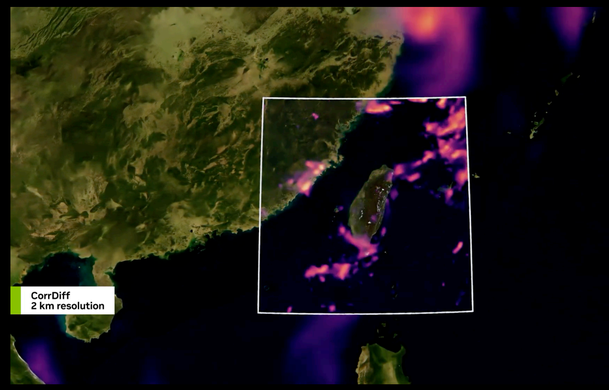
STEM Outreach

“Earth-2 Edu”

MILKYWAY3D.org

CorrDiff/Data-Conditioned Milky Way

“Galaxy-2”



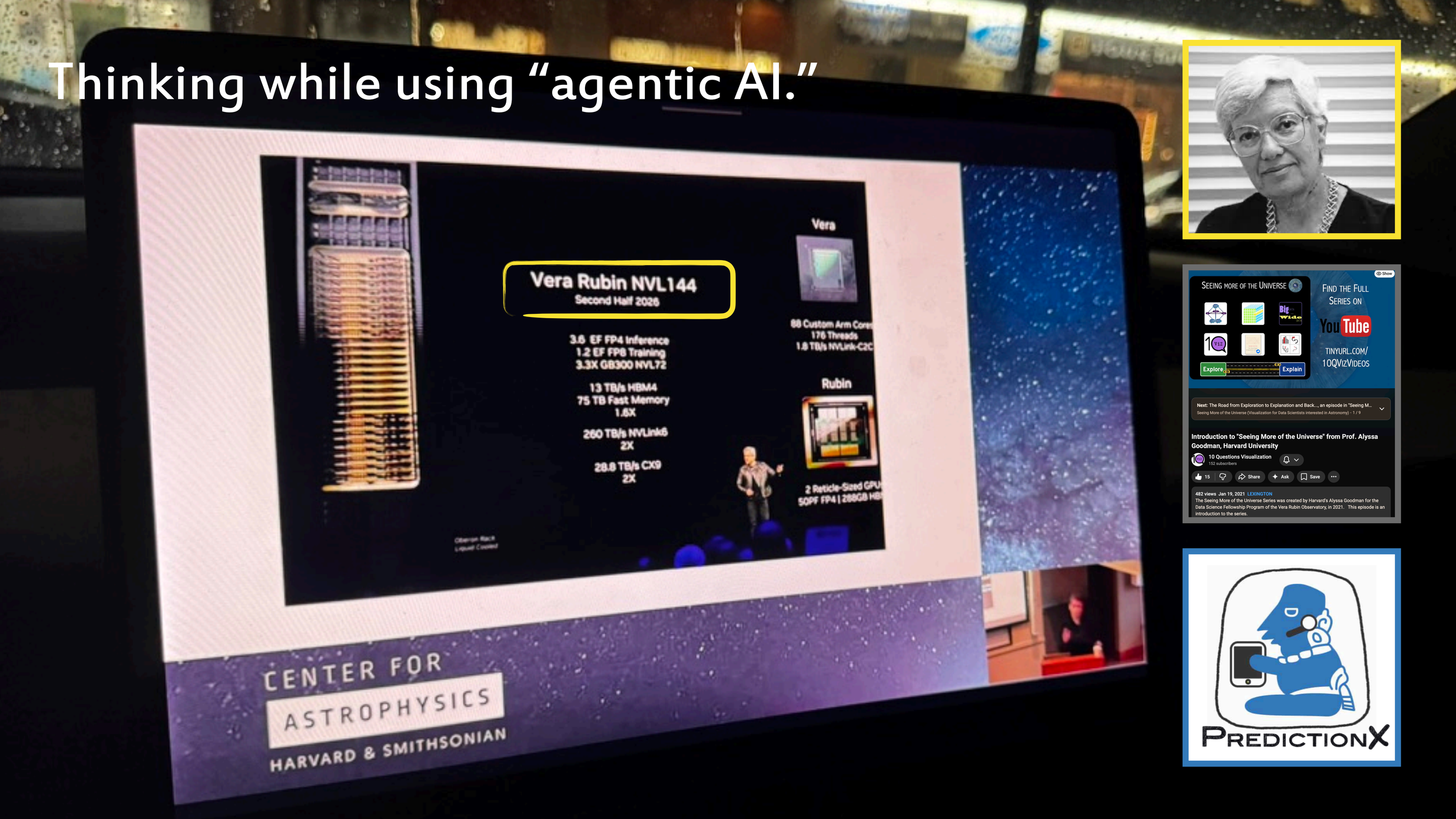
LIVE

glue
multidimensional data exploration

Exploratory Data Analysis

“LIVE Earth-2”

Thinking while using “agentic AI.”



The Washington Post
Democracy Dies in Darkness

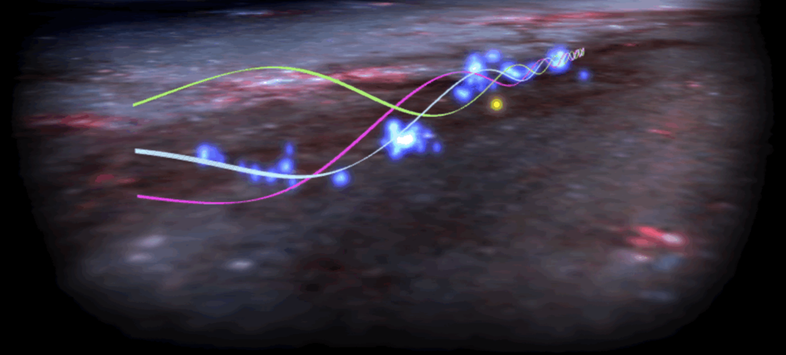
Science Space Animals Health Environment

Astronomers say mysterious galactic ‘wave’ may have once washed over Earth

Stretching across the night sky, a recently found chain of star-forming clouds is undulating through the galaxy

February 20, 2024 More than **1 year ago**

6 min 636



Radcliffe Wave Oscillation suggests no dark matter in Milky Way’s disk.

Article

The Radcliffe Wave is oscillating

https://doi.org/10.1038/s41586-024-07727-3
Received: 23 June 2023
Accepted: 26 January 2024
Published online: 20 February 2024
Check for updates

Ralf Konietzka[1,2✉], Alyssa A. Goodman[2], Catherine Zucker[3,4], Andreas Burkert[5,6], João Alves[6], Michael Foley[2], Cameron Swiggum[2], Maria Koller[6] & Núria Miret-Roig[6]

Our Sun lies within 300 parsecs of the 2.7-kiloparsecs-long sinusoidal chain of dense gas clouds known as the Radcliffe Wave[1]. The structure’s wave-like shape was discovered using three-dimensional dust mapping, but initial kinematic searches for oscillatory motion were inconclusive[2,3]. Here we present evidence that the Radcliffe Wave is oscillating through the Galactic plane while also drifting radially away from the Galactic Centre. We use measurements of line-of-sight velocity[4] for ^{12}CO and three-dimensional velocities of young stellar clusters to show that the most massive star-forming regions spatially associated with the Radcliffe Wave (including Orion, Cepheus, North America and Cygnus X) move as though they are part of an oscillating wave driven by the gravitational acceleration of the Galactic potential. By treating the Radcliffe Wave as a coherently oscillating structure, we can derive its motion independently of the local Galactic mass distribution, and directly measure local properties of the Galactic potential as well as the Sun’s vertical oscillation period. In addition, the measured drift of the Radcliffe Wave radially outwards from the Galactic Centre suggests that the cluster whose supernovae ultimately created today’s expanding Local Bubble[5] may have been born in the Radcliffe Wave.

In Fig. 1a (Supplementary Fig. 1), we present a three-dimensional (3D) map showing the most massive star-forming regions (including Orion, Cepheus, North America and Cygnus X) and embedded young stellar clusters associated with the Radcliffe Wave[1]. As expected, the young clusters and star-forming molecular clouds are co-located in three dimensions. The clusters were selected from a top-down view of the Galaxy, without an a priori selection criterion for height off the Galactic plane (Methods), suggesting that the clusters’ stars were born in the molecular clouds along the Radcliffe Wave. Therefore, these clusters’ motions can serve as a tracer of the Wave’s kinematics. Using the motion of clusters rather than of individual stars allows for more precise conclusions about kinematics, as the motion of individual stars within a host cluster are averaged out, reducing uncertainty.

The kinematic counterpart of the Radcliffe Wave’s spatial undulation is shown in Fig. 1b (Supplementary Fig. 1), where the z axis shows the vertical velocity of the stellar clusters, after accounting for the Sun’s motion[6]. On the basis of the sinusoid-like shape of the Radcliffe Wave, one could imagine it as either a travelling wave or a standing wave. Extended Data Fig. 1 and the animated version in Supplementary Fig. 4) shows how travelling and standing versions of the Radcliffe Wave would appear. For a travelling wave, the wave’s spatial shape fully determines vertical velocities, assuming a gravitational potential[7–10], as explained in Methods. For a standing wave, regions located at the zero crossings of the wave are at rest, and the velocities of regions located at spatial extrema are not constrained by the wave’s spatial structure. The fit shown as a solid black line in Fig. 1 compare Supplementary Fig. 1 and Methods) indicates that the Radcliffe Wave oscillates through the Galactic plane like a travelling wave, such that regions currently at the zero points of the wave (near the Galactic plane) move through at their greatest vertical velocity, while molecular clouds located at spatial extrema (farthest from the plane) are at their turning points, with zero vertical velocity. While a travelling wave provides an excellent fit to the observations, with a 90° phase offset between position (z) and velocity (v_z), as evident in Fig. 1 and in the interactive version in Supplementary Fig. 1), it also provides an extremum for the wave motion. When fitting a mixture model that allows for superpositions of a travelling wave and a standing wave, we find that a travelling wave is strongly preferred over a standing wave (Methods). A standing-wave model would have a zero phase offset between positional (z) and velocity (v_z) variations, which is inconsistent with the data (Methods).

For stars not far off the Galaxy’s disk compared with its scale height, a classical harmonic oscillator is often used to describe stars’ responses to the vertical Milky Way potential[10]. By analogy to a pendulum, however, when the amplitude of oscillation is large, one needs to consider nonlinear effects that are not accounted for in the formalism describing a purely harmonic oscillator. In the case of the Galaxy, these nonlinearities arise from the vertical decrease in the midplane density, becoming significant at vertical positions on the order of the Galaxy’s vertical scale height. Therefore, in describing the motion of a pendulum oscillating far from vertical, we model the wave’s motion within the Galactic potential[9–11] as an anharmonic oscillator. In the modelling, we account for the Galaxy’s midplane density declining with height and also with galactocentric radius (Methods). Treating the Radcliffe Wave as a single coherent structure in space and velocity, responding to the Galactic potential, the structure is well modelled as a damped sinusoidal wave with a maximum amplitude of about 220 pc and a mean wavelength of about 2 kpc. The corresponding maximum vertical velocity is about 14 km s^{-1}.

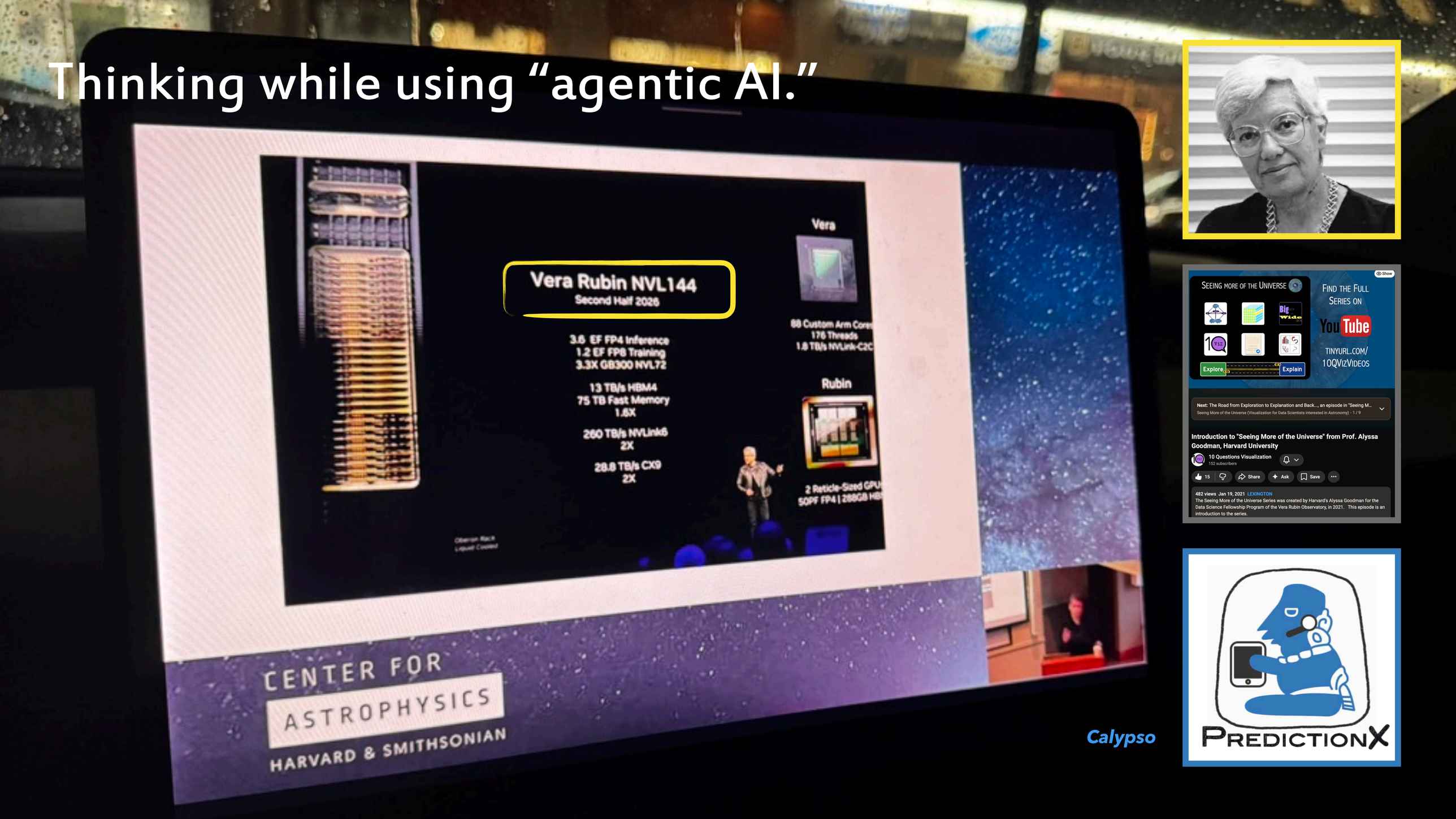
[1]Harvard-Smithsonian Center for Astrophysics | Harvard & Smithsonian, Cambridge, MA, USA. [2]Ludwig-Maximilians-Universität München, Munich, Germany. [3]Max Planck Institute for Extraterrestrial Physics, Garching, Germany. [4]Space Telescope Science Institute, Baltimore, MD, USA. [5]University Observatory Munich, Munich, Germany. [6]Department of Astrophysics, University of Vienna, Vienna, Austria. ✉e-mail: ralf.konietzka@cfa.harvard.edu

The Radcliffe Wave is oscillating

The spatially and kinematically coherent Radcliffe Wave serves as a unique envir
obtaining insights into local Galactic dynamics. As explained above, the kinema
signature of the Radcliffe Wave can be described in a self-consistent manner, usi
response to the local Galactic potential. Near the Sun, this potential is dominate
gravitational attraction of baryonic matter in the form of stars and gas[21]. In addi
matter, which is known to envelope galaxies in a spherical ‘dark halo’ is also expe
the local disk kinematics[22]. In Methods, we use the coherent motion of the Radc
assuming that it is controlled by the combined gravity of baryonic and dark mat
estimate the properties of the local Galactic potential. We find a total midplane d
$0.085^{+0.021}_{-0.017}\ M_\odot\ \mathrm{pc}^{-3}$, consistent with conventional approaches to deriving properties of the local Galactic mass distribution[21,23,24,25,26]. By incorporating direct observations of the gaseous and stellar-mass components[21], and thereby constraining the amount of the baryonic mass component in the solar neighbourhood, we also infer the local amount of dark matter (Methods). Our analysis yields a local density of $0.001^{+0.024}_{-0.021}\ M_\odot\ \mathrm{pc}^{-3}$ for a spherical dark halo. If dissipative dark matter exists, it is possible that dark matter may also accumulate in the form of a very thin disk relative to the baryonic mass component[27]. We find an upper bound for a hypothetical dark disk of $0.1^{+3.3}_{-3.3}\ M_\odot\ \mathrm{pc}^{-2}$. Simultaneously, we also utilize the Radcliffe Wave to derive the frequency of the Sun’s oscillation through the Galactic plane (Methods). We find that the Sun oscillates through the disk of the Milky Way with a period of $95^{+12}_{-10}\ \mathrm{Myr}$, implying that our Solar System crosses the Galactic Disk every $48^{+6}_{-5}\ \mathrm{Myr}$, consistent with standard values[14] (Methods).

Konietzka et al. 2024, *Nature*

Thinking while using “agentic AI.”



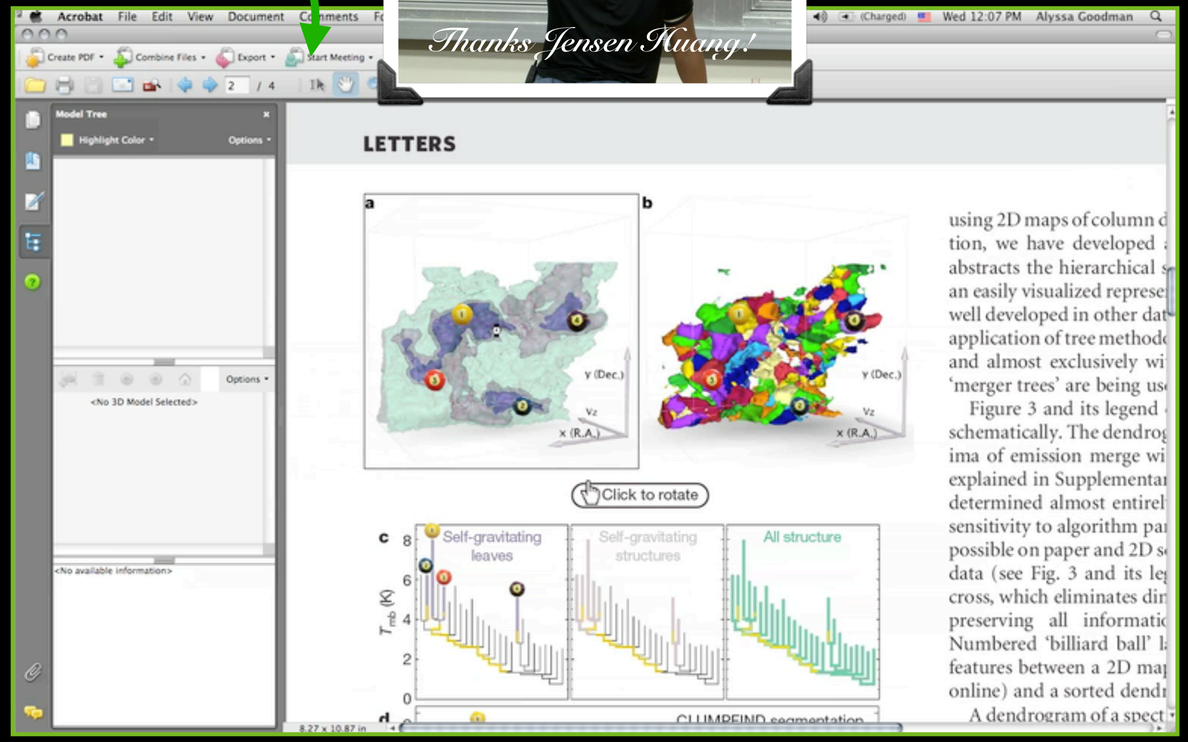
Calypso

PREDICTIONX

Exploratory Data Analysis, Then & Now

NVIDIA

2009

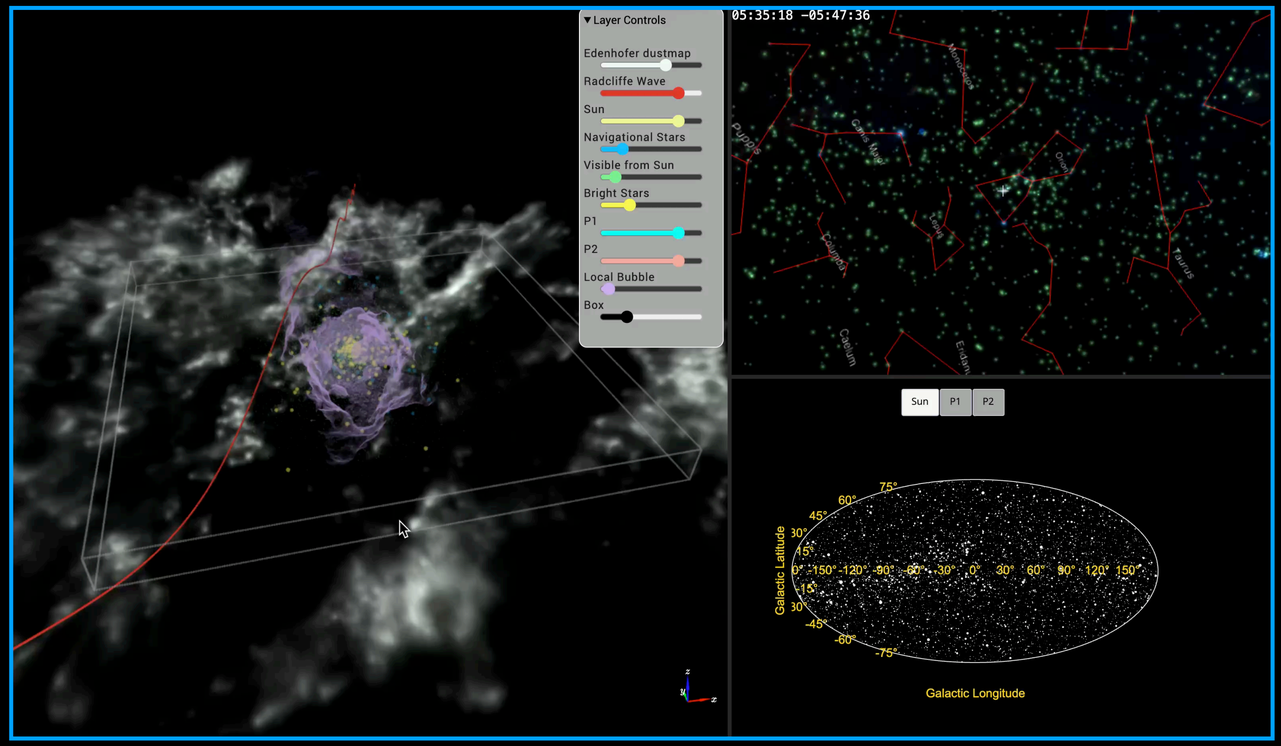


Goodman et al. *Nature*, 2009

AstronomicalMedicine@

3D PDF

TODAY



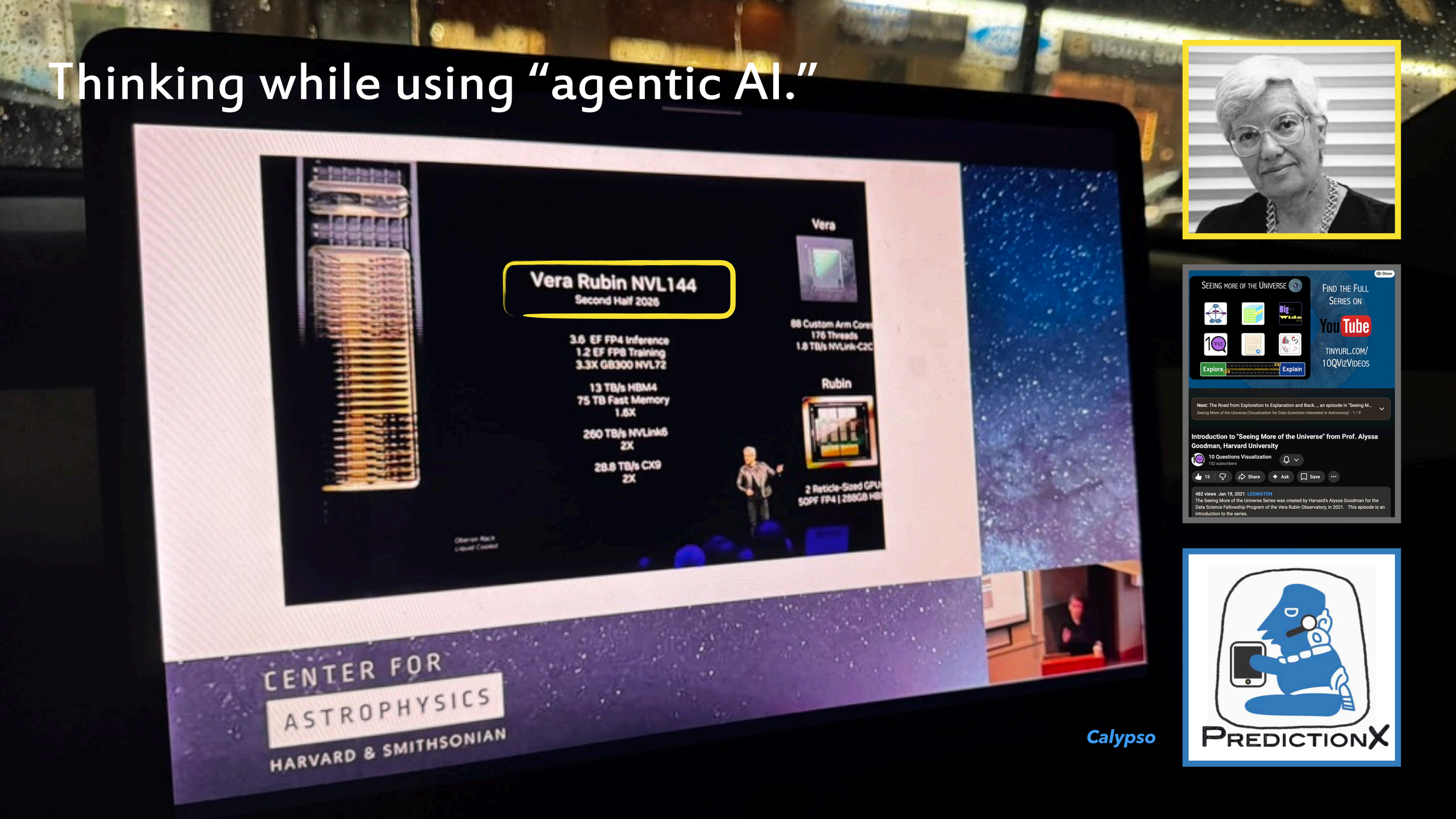
Klessen et al. *(to be submitted to Nature* 2025)

glue

LIVE

in-browser

Thinking while using “agentic AI.”



Calypso

PREDICTIONX

The Prediction Project

PredictionX: The Past & Present of the Future

Essentials

Predictive Systems Framework

The "Padua Rainbow"

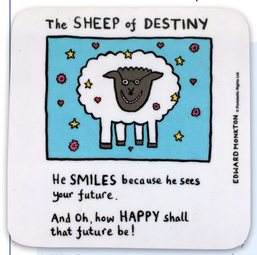
Understanding Uncertainty

Study Design

Timelines

Why predict?

Omens, Oracles & Prophecies



Egyptian Priests

Tarot

The Diviner's Guide

Chinese Oracle Bones

Oracle of Delphi

Aztec Rituals

Turkish Tasseography

Maya Spacetime

Yoruba Ifa

Casting Lots

Greek Astronomy

Astrology

Comets of Doom

cross-cultural conversations

The Rise of Theory

Ancient Mesopotamia, Egypt, Greece & Rome

Islamic Science

The Path to Newton

Indian Mathematics

European Renaissance

Lost without Longitude (Navigation

Help, I'm Lost

Tools of the Navigator

The Royal Society

Modern Prediction

John Snow & Cholera

Cholera Map

Health

Epidemiology

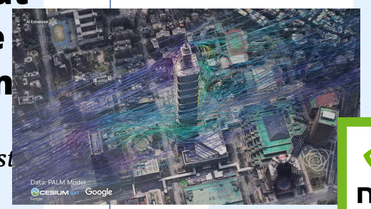
Personal Genomics

Population Genetics

Wealth

Climate & Wealth

Behavioral Economics



The Future of the Future

AI, Derek's Day

Philosophy

ertainty

Earth

Climate & Energy

Climate Policy

Tent Tarot

Earthquakes

Space

Futures of our Universe

SETI

Interactive Resource ▶video(s) Coming Soon visit predictionx.org for more information on the Prediction Project

edX HarvardX predictionx.org

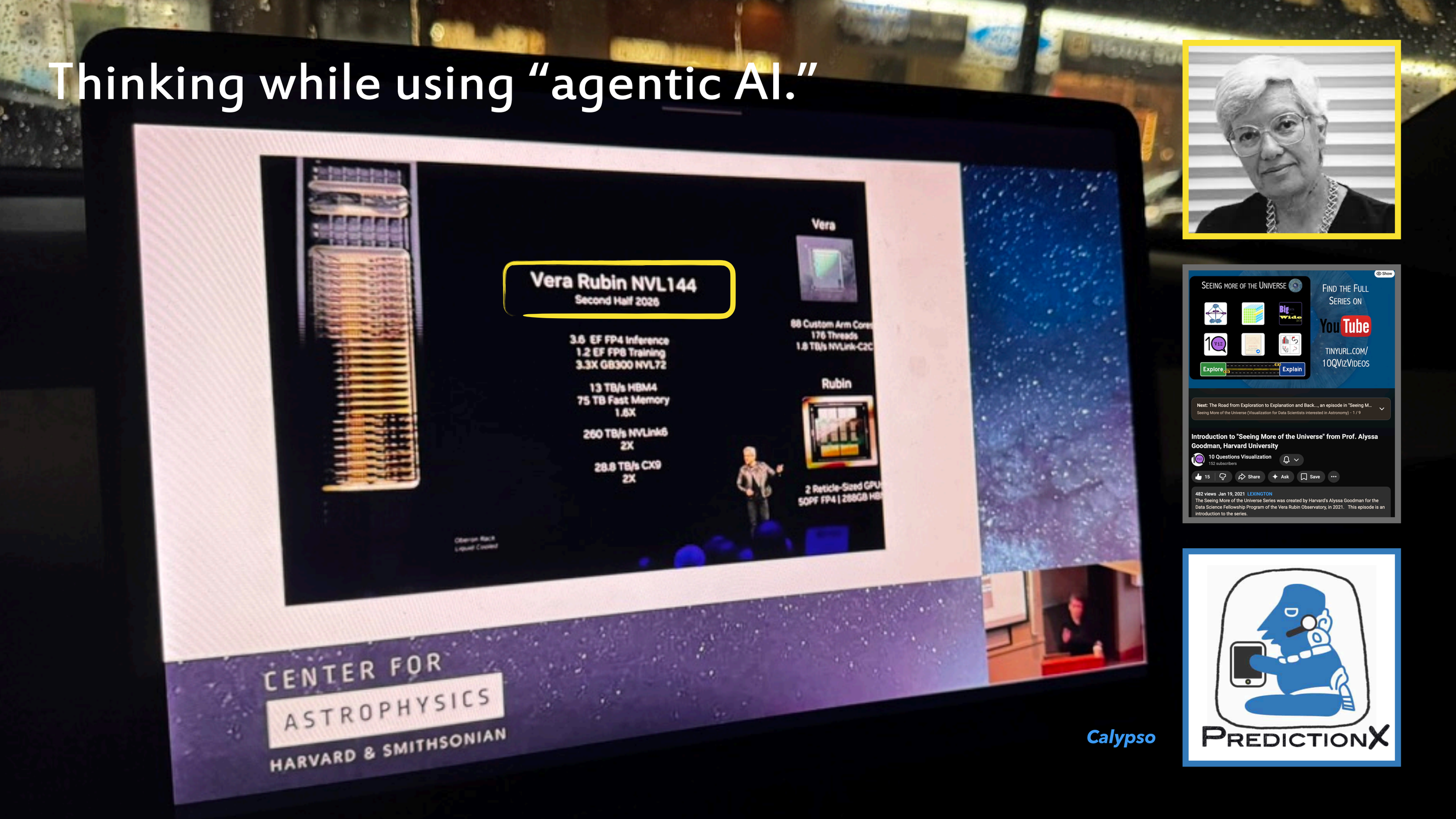
Harvard College Program in General Education
Explore. Expand. Engage.

LabXchange™ Science and learning-connected.

The Future of the Future

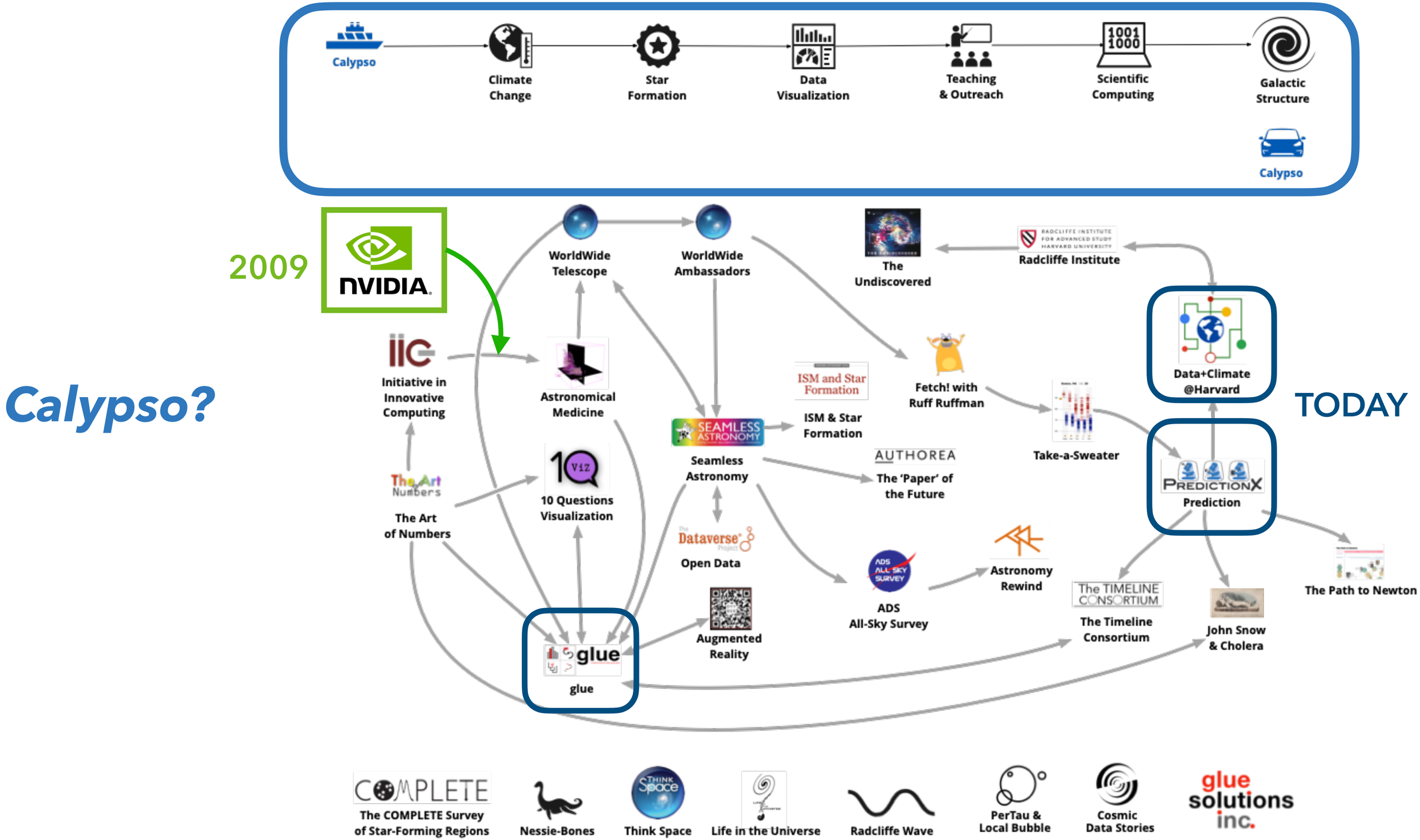
20th century

Thinking while using “agentic AI.”



Calypso

PREDICTIONX


Calypso
Climate Change
Star Formation
Data Visualization
Teaching & Outreach
Scientific Computing
Galactic Structure
Calypso
2009
NVIDIA
Calypso?
WorldWide Telescope
WorldWide Ambassadors
The Undiscovered
Radcliffe Institute
Initiative in Innovative Computing
Astronomical Medicine
ISM and Star Formation
ISM & Star Formation
Fetch! with Ruff Ruffman
Take-a-Sweater
Data+Climate @Harvard
TODAY
Seamless Astronomy
AUTHOREA
The 'Paper' of the Future
Prediction
The Art of Numbers
10 Questions Visualization
Dataverse
Open Data
ADS All-Sky Survey
Astronomy Rewind
The Timeline Consortium
John Snow & Cholera
The Path to Newton
Augmented Reality
glue
The COMPLETE Survey of Star-Forming Regions
Nessie-Bones
Think Space
Life in the Universe
Radcliffe Wave
PerTau & Local Bubble
Cosmic Data Stories
glue solutions inc.

What could this mean for NVIDIA + PredictionX/MilkyWay3D/LIVE?

CosmicDS

PREDICTIONX

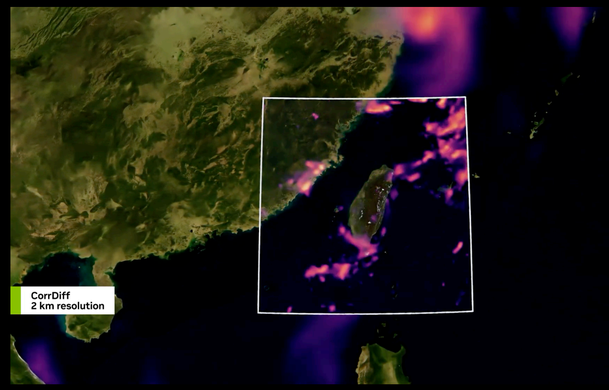
STEM Outreach

“Earth-2 Edu”

MILKYWAY3D.org

CorrDiff/Data-Conditioned Milky Way

“Galaxy-2”



LIVE

glue
multidimensional data exploration

Exploratory Data Analysis

“LIVE Earth-2”

glue an open-source software environment for exploratory data analysis (EDA)

multidimensional data exploration

multiple data sets analyzed together

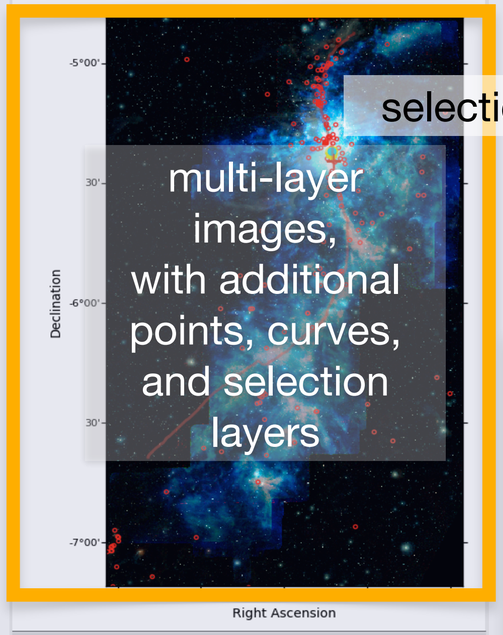
selections across data sets

data sets attributes linked (UI not shown)

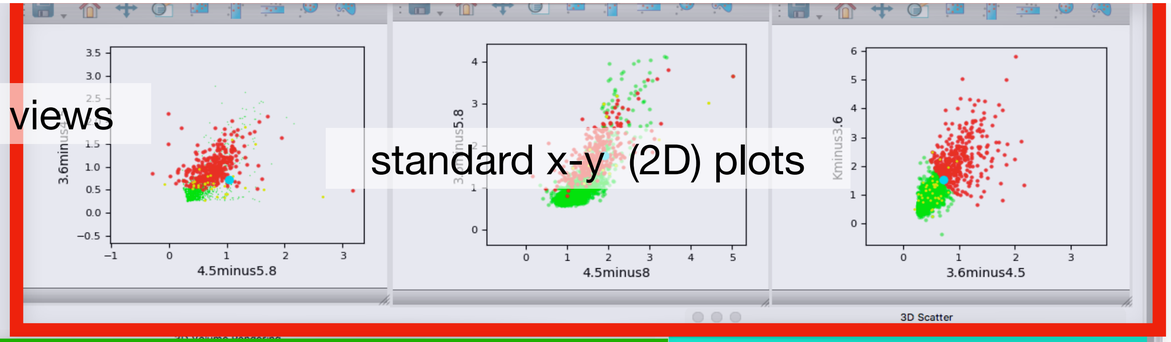
table viewer (not shown)

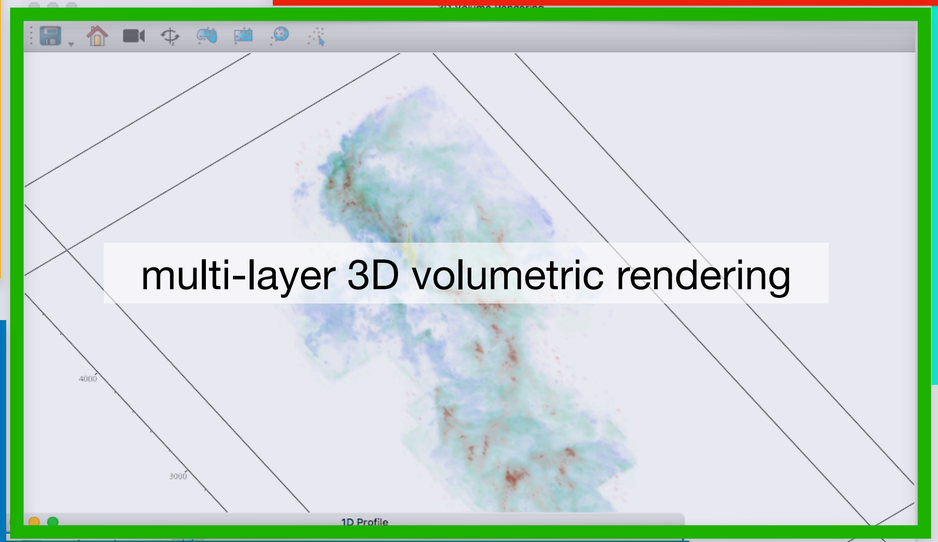
stats calculator (not shown)

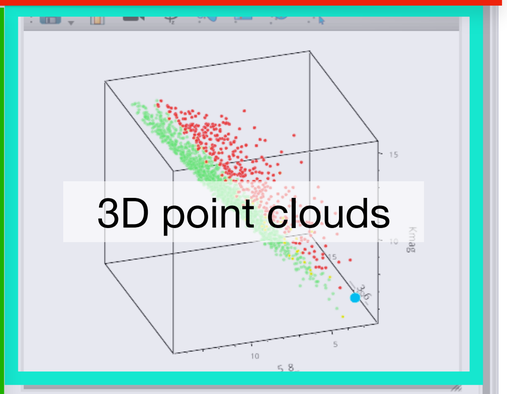
custom plot types (not shown)

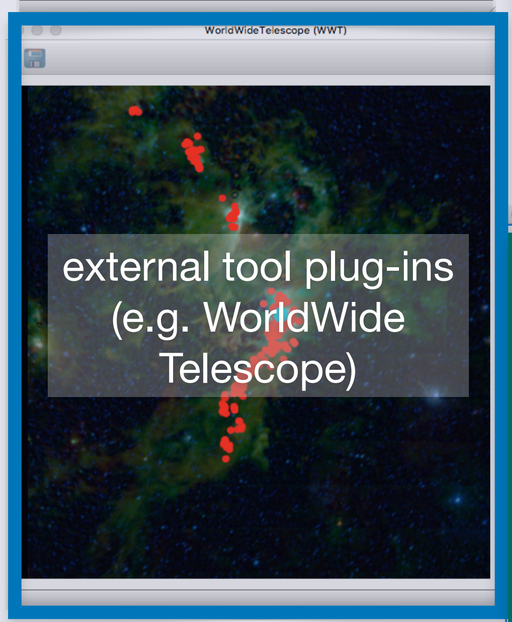


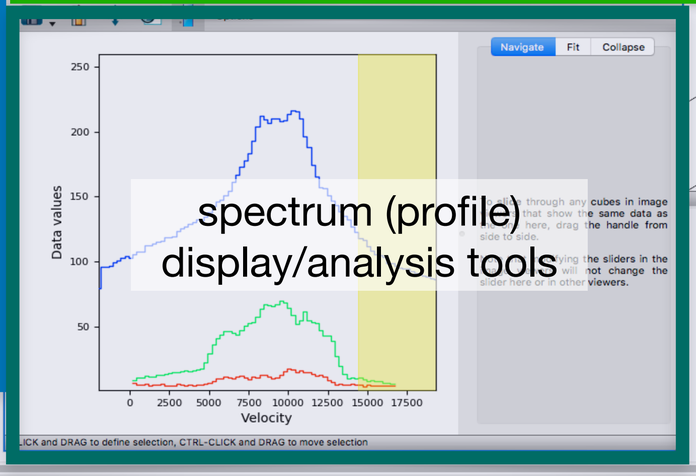
selections across all views

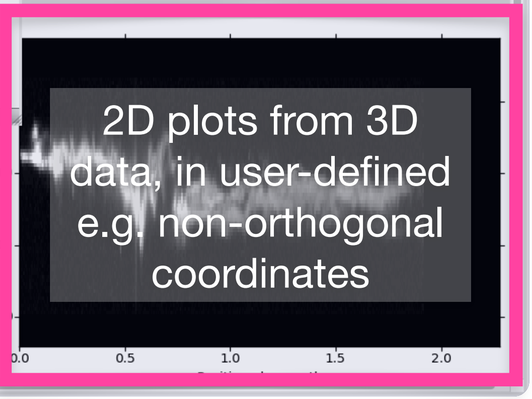










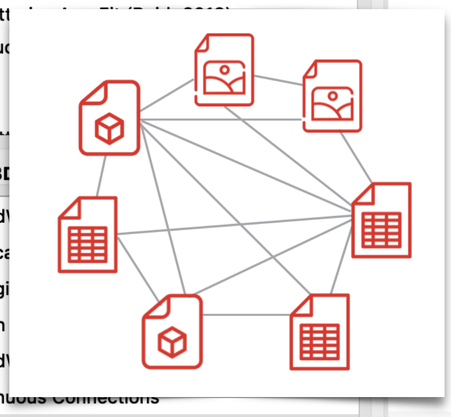


Data Collection

glue
multidimensional data exploration

Green 2019 3D Dust
Local Arm Fit (Reid+2016)
Major Cloud Catalog
Maser Catalog (Reid+2014,2016)
Sagitt
Tenuc
Sun

Subsets



Plot Layers - 3D

Rad
Loca
Sagi
Sun
Rad
Tenuous Connections

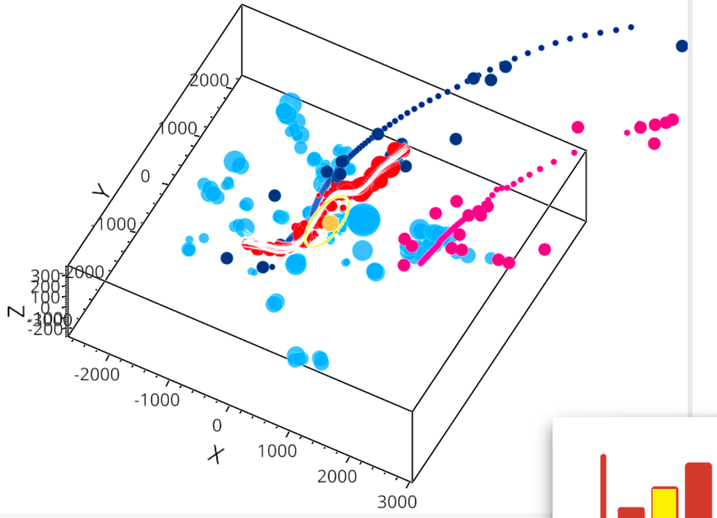
Attribute: PRIMARY
Limits: 1 5
Color:

Plot Options - 3D Volume Rendering

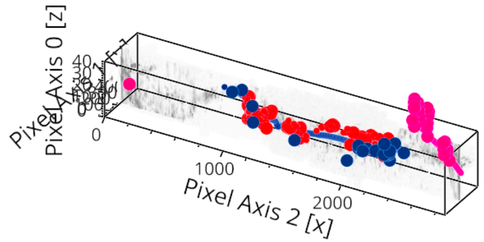
x axis: Pixel Axis 2 [x]
min/max: 38.2241 ⇄ 1160.78
stretch: 1.00
y axis: Pixel Axis 1 [y]
min/max: 38.2241 ⇄ 1160.78
stretch: 1.00
z axis: Pixel Axis 0 [z]
min/max: 5.95402 ⇄ 193.046
stretch: 1.00
reference: Green 2019 3D Dust

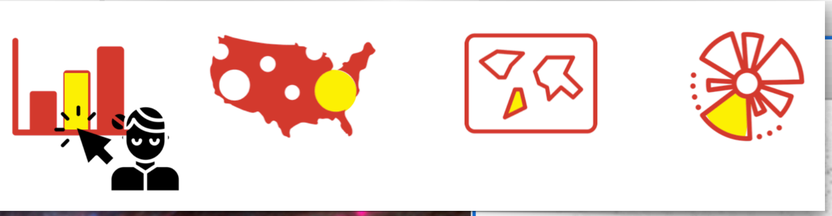
Tab 1

3D Scatter



3D Volume Rendering

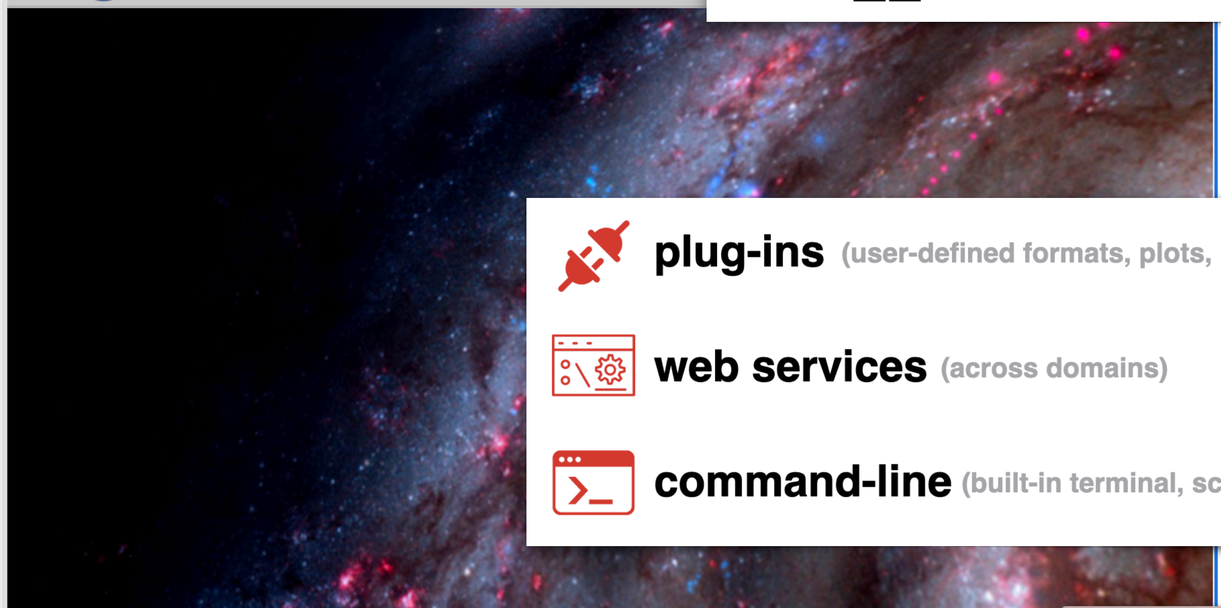


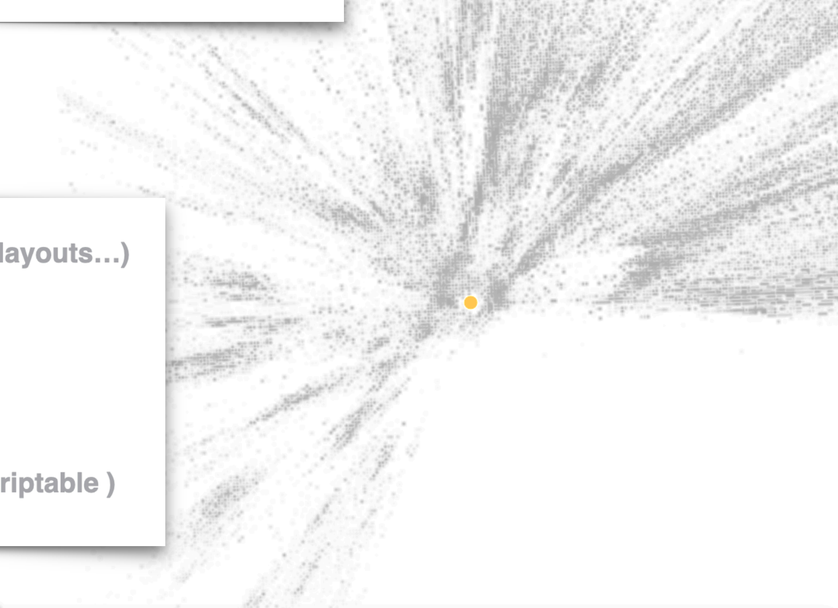


3D Volume Rendering

WorldWi

WORLDWIDE TELESCOPE





plug-ins (user-defined formats, plots, layouts…)

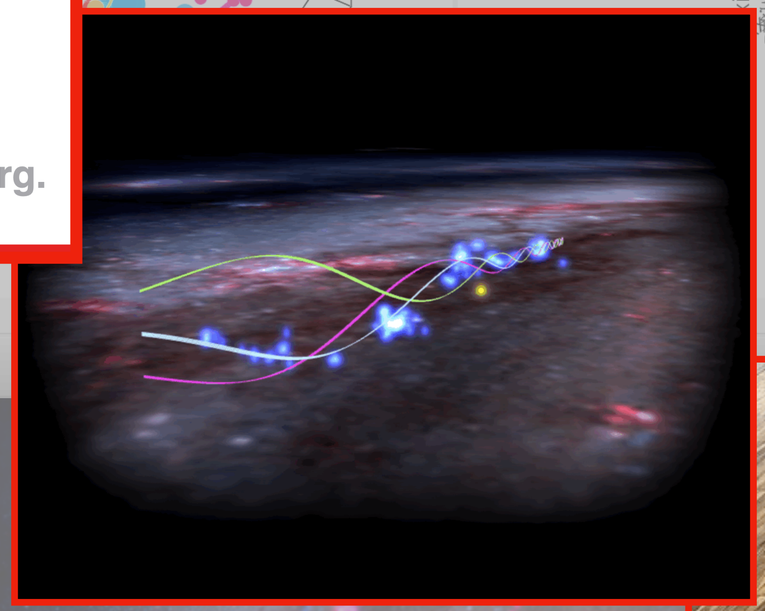
web services (across domains)

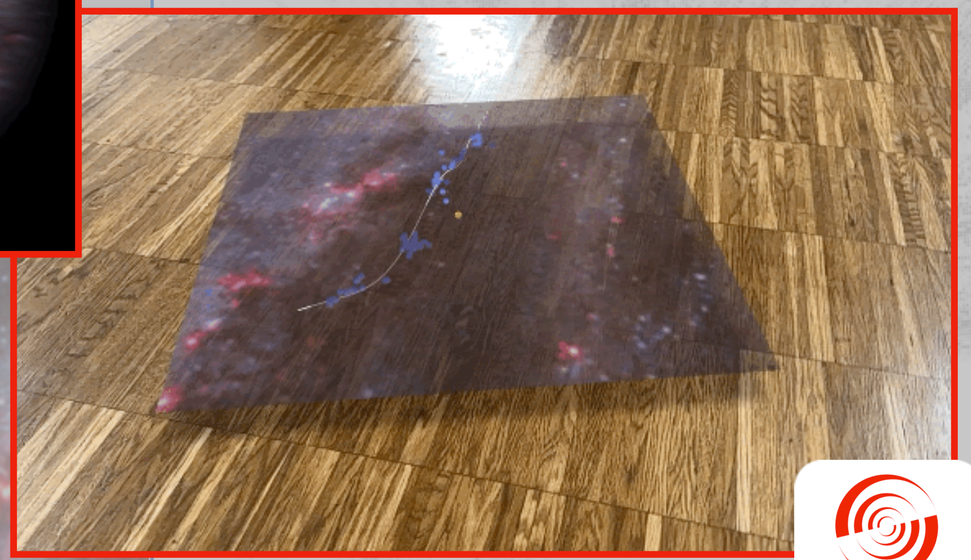
command-line (built-in terminal, scriptable)

BONUS: save, share, or publish what you learn—

- save "sessions" to continue where you left off
- export graphics
- use/export to Jupyter environments
- export to plot.ly (javascript)
- export to augmented reality

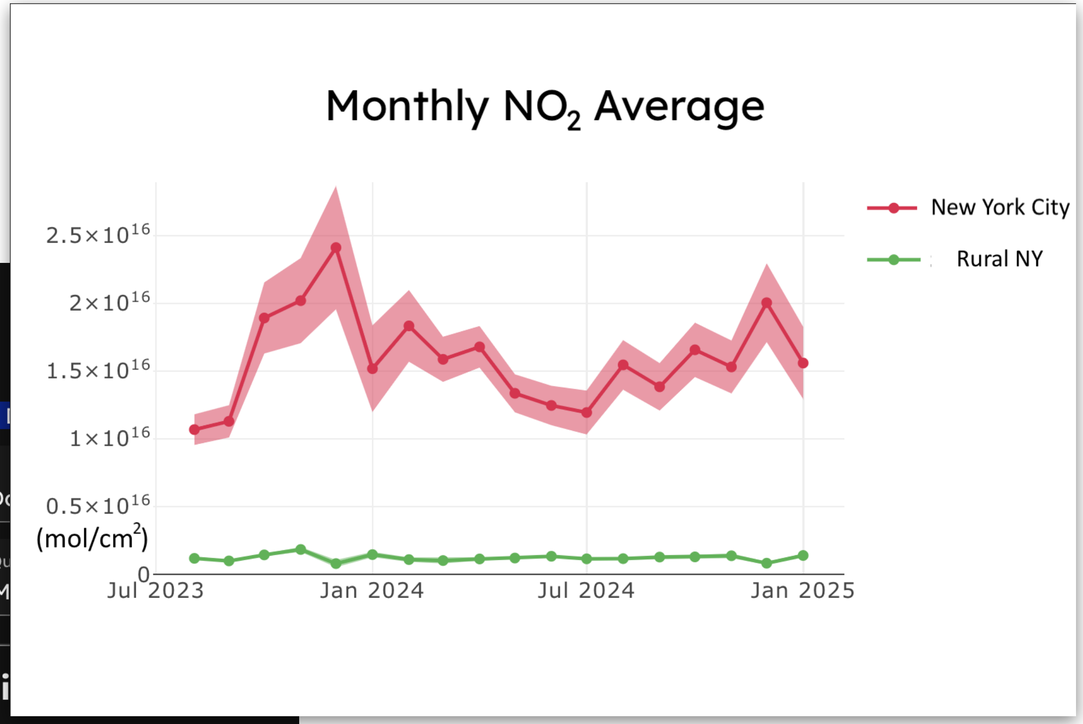
learn how at glueviz.org.

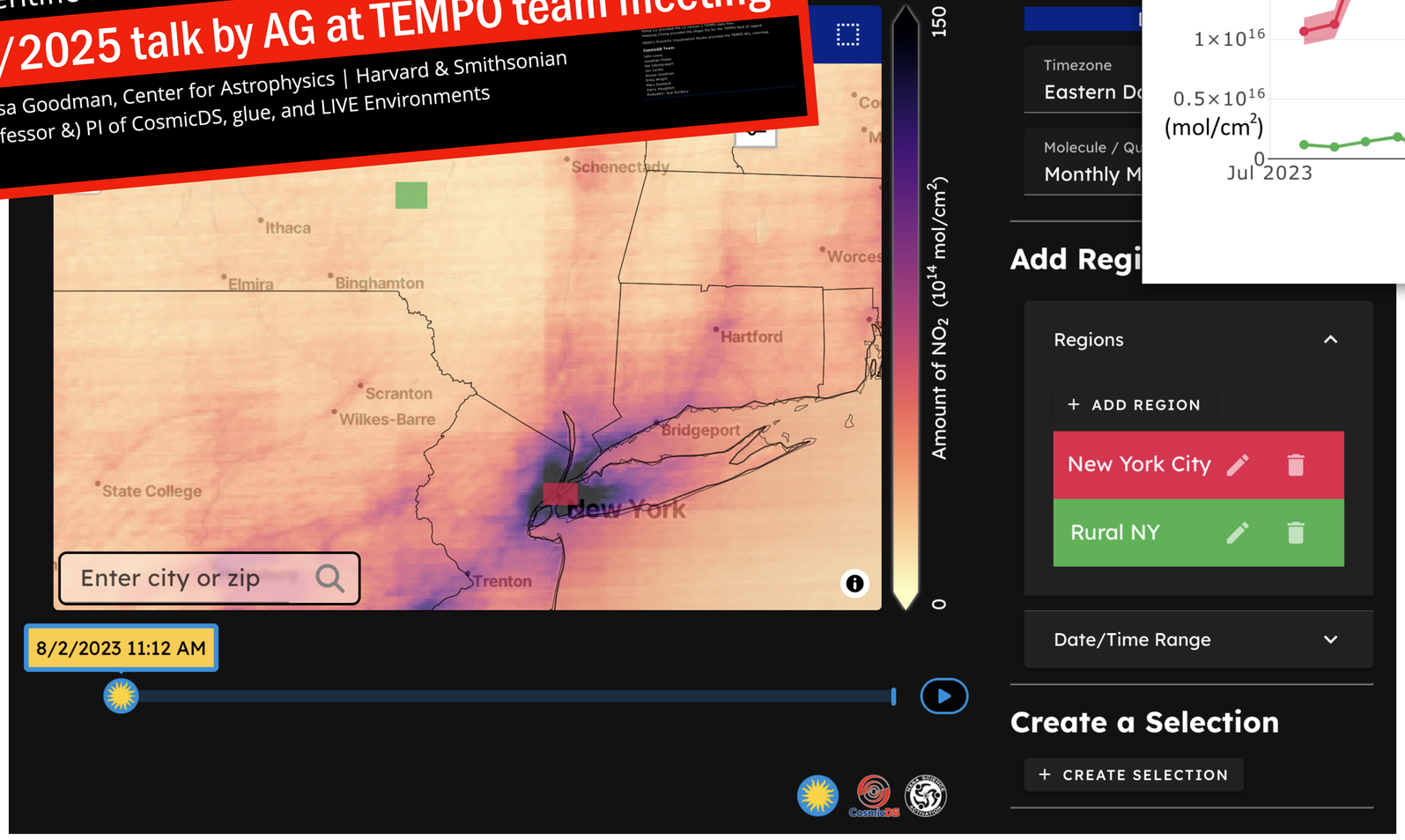




CosmicDS

nalysis





glue multidimensional data exploration

LIVE

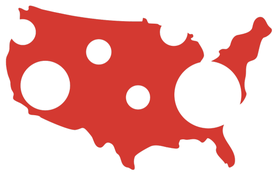
CosmicDS

TEMPO

“graphs”

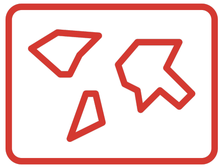
common statistical graphics

(scatterplots, histograms, tables, curves, overlays)



maps & images

(greyscale, color, contours, layer control…)

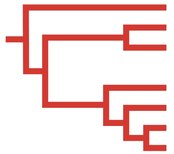


3D displays

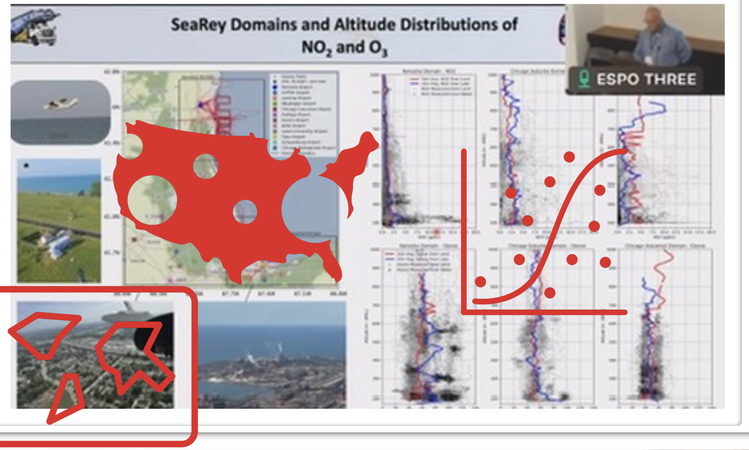
(scatter plots, volumetric rendering, sliders…)

specialized & custom charts

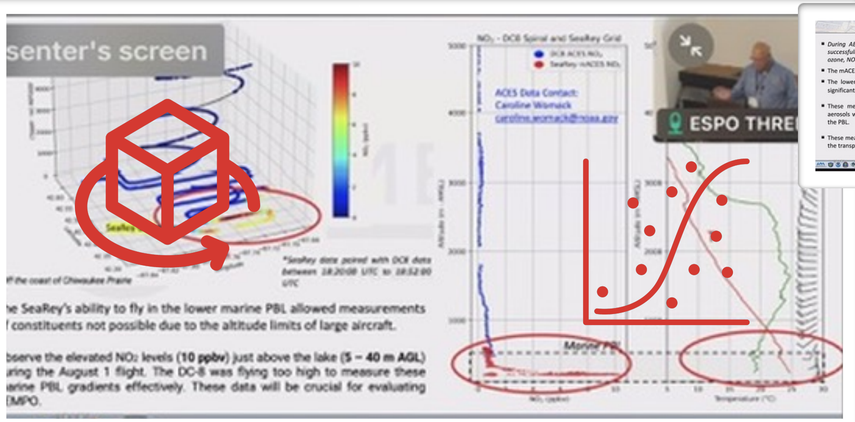
(dendrograms, polar plots, + domain-specific options)

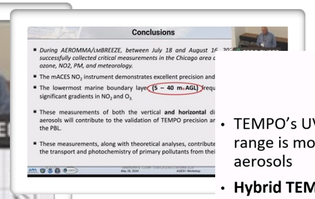


SeaRey Domains and Altitude Distributions of NO_2 and O_3



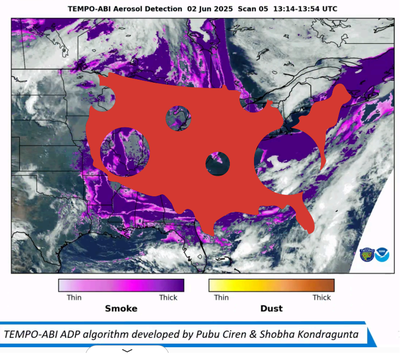
senter's screen

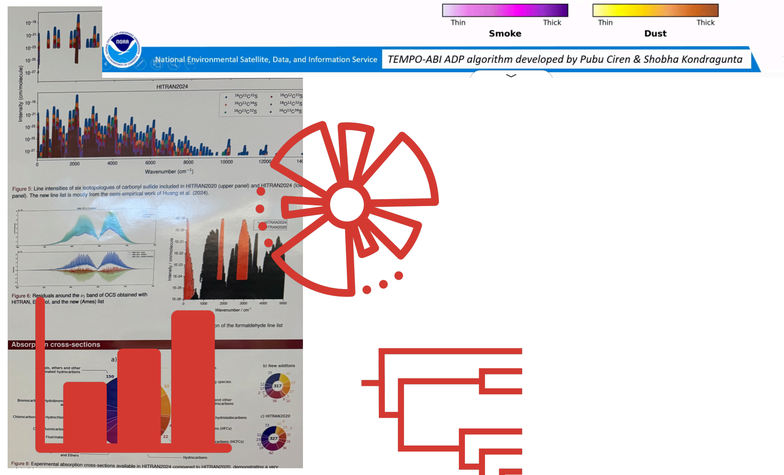




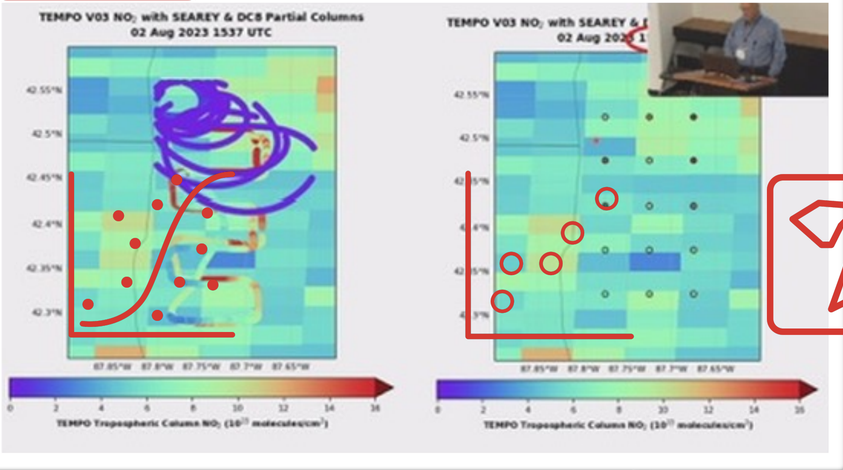
TEMPO-ABI Synergy Enhances Smoke Detection

- TEMPO's UV-Deep Blue spectral range is more sensitive to absorbing aerosols
- **Hybrid TEMPO-ABI Aerosol Detection** algorithm takes advantage of the synergy between the two sensors
- Much greater coverage of the wildfire smoke than using ABI alone!
- Also provides info on the relative intensity of smoke/dust (thick vs thin)

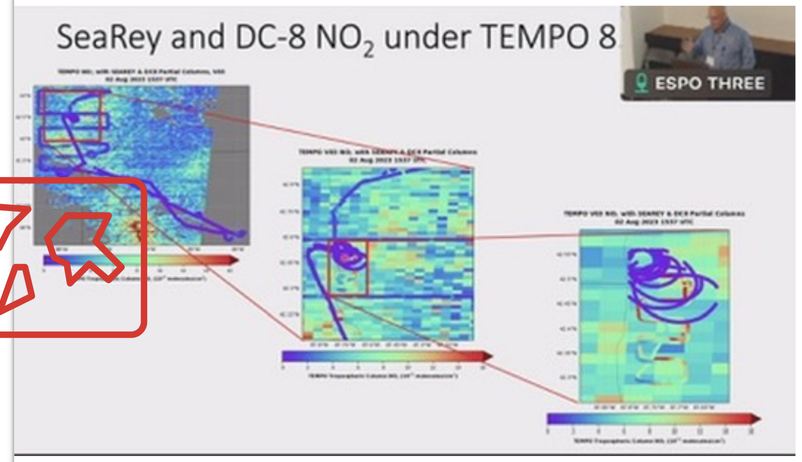


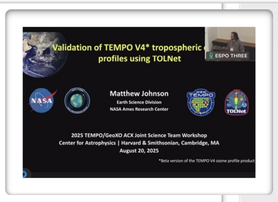


TEMPO V03 NO_2 with SEAREY & DC8 Partial Columns 02 Aug 2023 1537 UTC

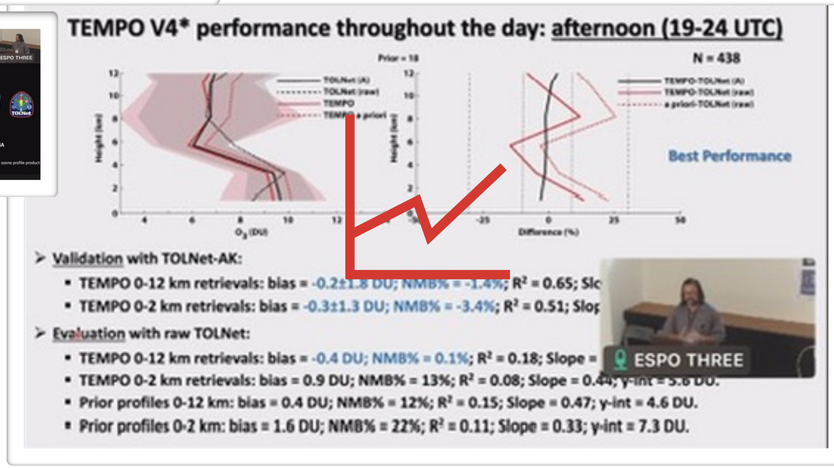


SeaRey and DC-8 NO_2 under TEMPO 8



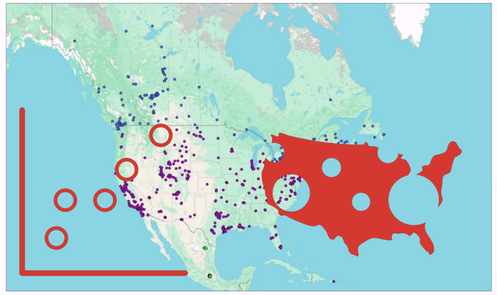


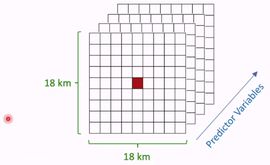
TEMPO V4* performance throughout the day: afternoon (19-24 UTC)



- Validation with TOLNet-AK:
  - TEMPO 0-12 km retrievals: bias = -0.2±1.8 DU; NMB% = -1.4%; R^2 = 0.65; Slo
  - TEMPO 0-2 km retrievals: bias = -0.3±1.3 DU; NMB% = -3.4%; R^2 = 0.51; Slog
- Evaluation with raw TOLNet:
  - TEMPO 0-12 km retrievals: bias = -0.4 DU; NMB% = 0.1%; R^2 = 0.18; Slope =
  - TEMPO 0-2 km retrievals: bias = 0.9 DU; NMB% = 13%; R^2 = 0.08; Slope = 0.44; y-int = 5.6 DU.
  - Prior profiles 0-12 km: bias = 0.4 DU; NMB% = 12%; R^2 = 0.15; Slope = 0.47; y-int = 4.6 DU.
  - Prior profiles 0-2 km: bias = 1.6 DU; NMB% = 22%; R^2 = 0.11; Slope = 0.33; y-int = 7.3 DU.

Using Deep Learning and TEMPO to Estimate Surface NO_2





Predictor Variables:
TEMPO TOT
Meteorolog
MODIS EVI
Urban-relat
Surface elev

Ten-Fold Cross-Validation Results

	R	IOA	MAB (%)	MB (%)	RMSE (%)
2023	0.98	0.98	0.76 (10)	-0.01 (-0.2)	1.33 (18)

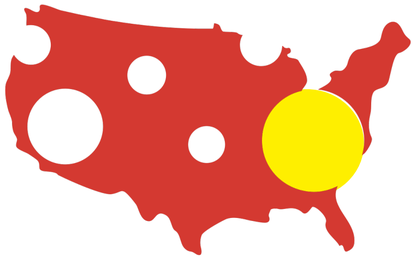
MAB: Mean Absolute Bias (ppb)
MB: Mean Bias (ppb)
RMSE: Root Mean Square Error (ppb)

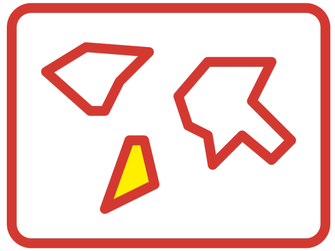
Spatial Cross-Validation Results

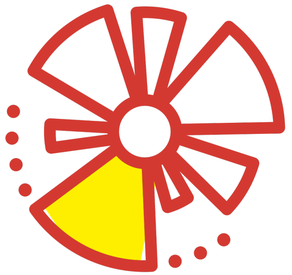
	R	IOA	MAB (%)	MB (%)	RMSE (%)
2023	0.90	0.94	1.4 (25)	0.2 (6)	2.0 (35)

MAB: Mean Absolute Bias (ppb)
MB: Mean Bias (ppb)
RMSE: Root Mean Square Error (ppb)

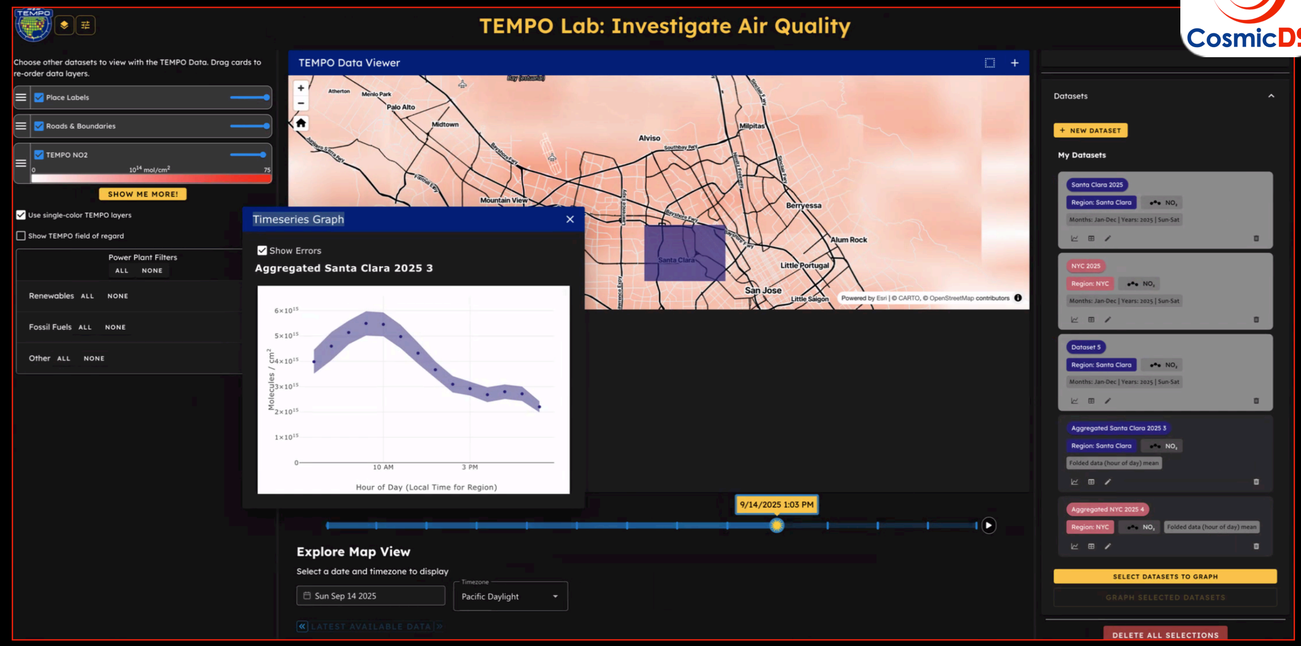
selections propagate across all graphs







Exploratory Data Analysis



"Interactive Visualization of High-Resolution, Global-Scale Climate Data in the Cloud"

LIVE = glue + jupyter + +++

LIVE

an open-source exploratory data analysis (EDA) software coordination project, led by Alyssa Goodman (glue) & Fernando Pérez (Jupyter), with MANY collaborators

LIVE-Env.org

NSF POSE Phase II proposal submitted 9/2/2025. Read the Proposal & More

LIVE Environments | About LIVE | Capabilities | Tech | Who are we? | Learn More

Linkable Interactive Visualization Exploration (LIVE) Environments

What is LIVE?

"MilkyWay3D.org" is the "Demonstration Project" for LIVE-Astro

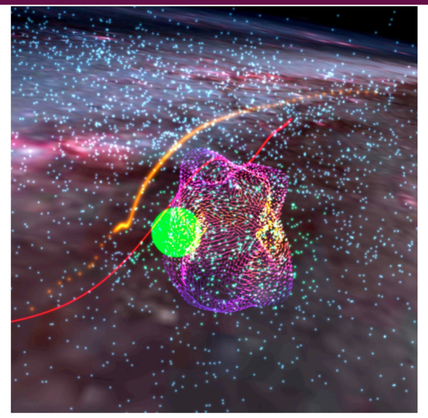
MILKYWAY3D.org

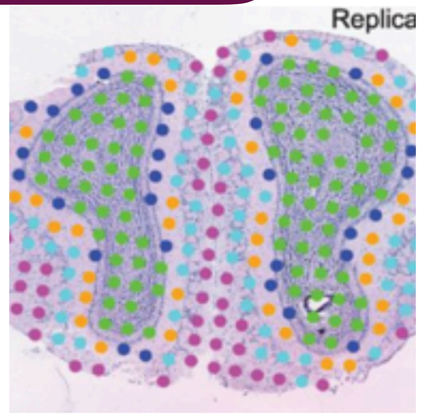
nments.

visualization challenges across astronomy

merging open-source ecosystem. LIVE

cts to refine and ensure the platform's utility,





LIVE Astro | LIVE Bio | LIVE GIS

Software (red outline= used as part of LIVE today)	Function	Browser (B) Desktop (D) Python anywhere (P)	(Filter by) Usage	Priority Level	Description	Level of Connection to LIVE as of 2025 1=working 2=prototype 3=discuss'ns underway 4=not yet	GitHub Link	Coordinatiing Contacts
GeoJupyter					collaborative cloud based tool to analyze and visualize geospatial data in JuypterLab.	2	github.com/geojupyter/	Pangeo & Jupyter
JupyterGIS					geospatial extensions for Jupyter notebooks	2	github.com/geojupyter/ju	GeoJupyter / Pangeo
Leaflet					Interactive maps for web applications	1	github.com/Leaflet	LeafletJS community
GeoPandas	Data Analysis	Py	🌎	2	Geospatial data analysis with pandas	3	github.com/geopandas/	GeoPandas community / NumFOCUS
Cryocloud	Community Platform	B	🌎	3	Cloud platform for cryosphere research	4	github.com/CryoInTheClou	NASA Cryosphere teams
Pangeo	Community Platform	B	🌎	3	Scalable geoscience data analysis platform, ecosystem built around Dask, xarray, Jupyter	4	github.com/pangeo-data/	Pangeo community
OpenSpace	Earth & Sky Viewer	D	🪶🌎	1	3D space visualization platform (used w/glue and WWT in MilkyWay3D.org)	1	github.com/OpenSpace	American Museum of Natural History/NASA
WorldWide Telescope (WWT)	Earth & Sky Viewer	D, B	🪶🌎	1	displays all-Sky and whole-Earth images with inset image and data overlay options	1	github.com/WorldWideTel	SAO/Harvard
					2-D visualization system for Jupyter, based on			Jupyter

LIVE

Connect **Omniverse**?
Make **Earth-2** the **LIVE-GIS** demonstration project?

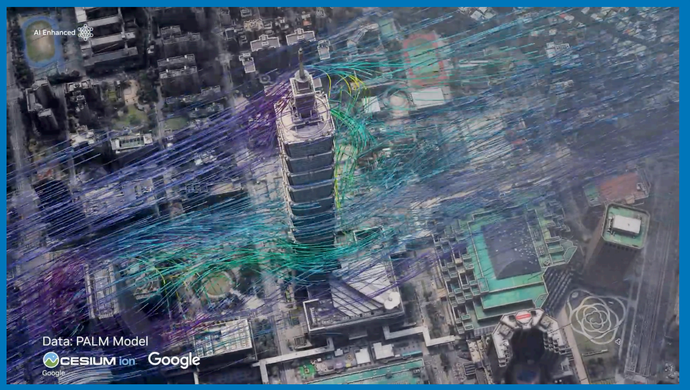
What could this mean for NVIDIA + PredictionX/MilkyWay3D/LIVE?

CosmicDS

PREDICTIONX

STEM Outreach

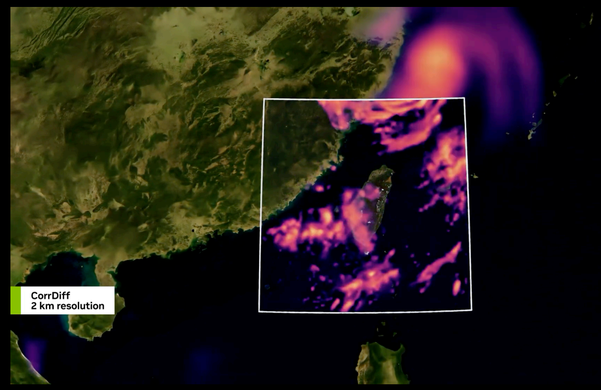
“Earth-2 Edu”



MILKYWAY3D.org

CorrDiff/Data-Conditioned Milky Way

“Galaxy-2”

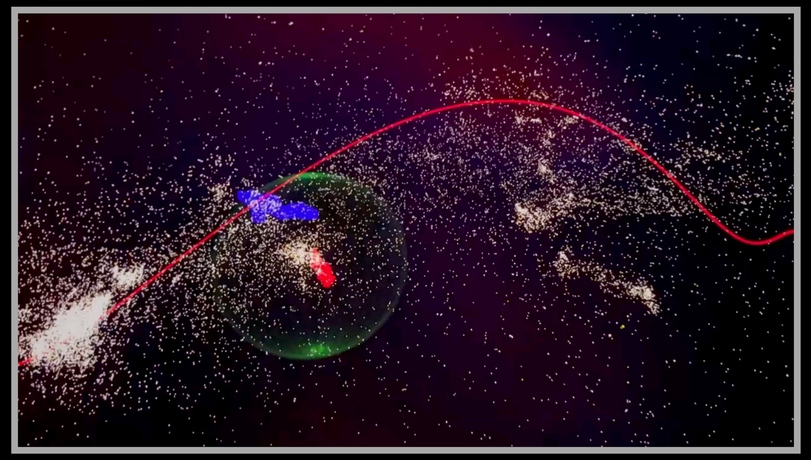


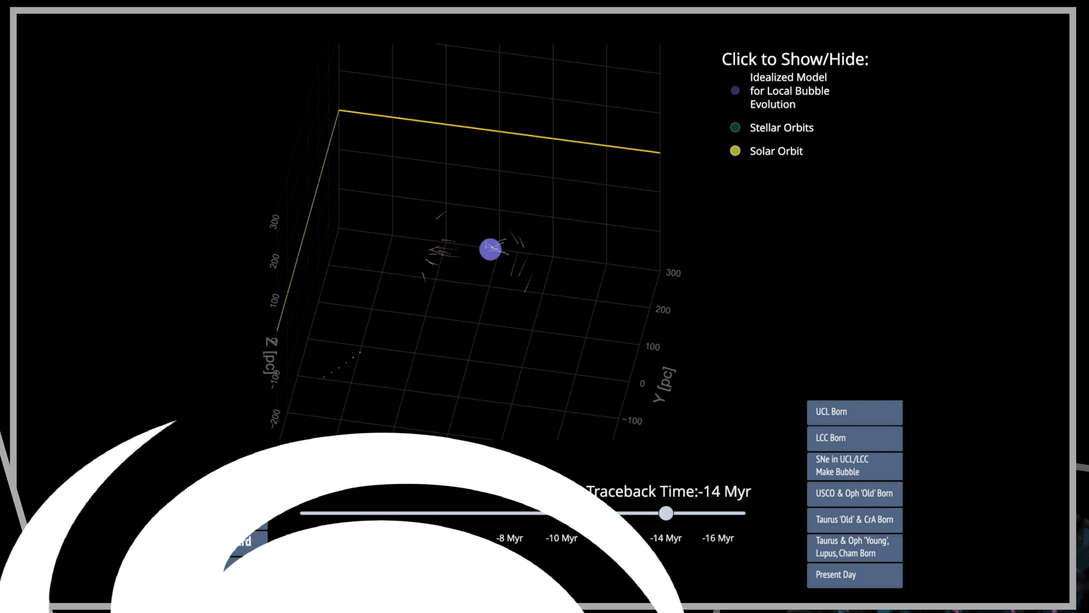
LIVE

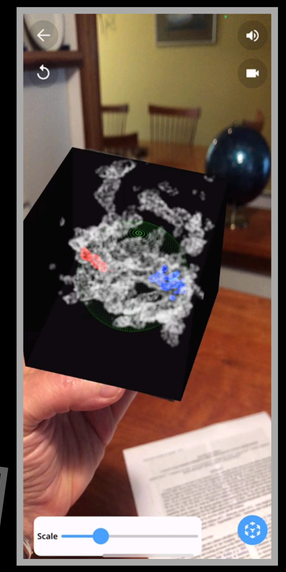
glue
multidimensional data exploration

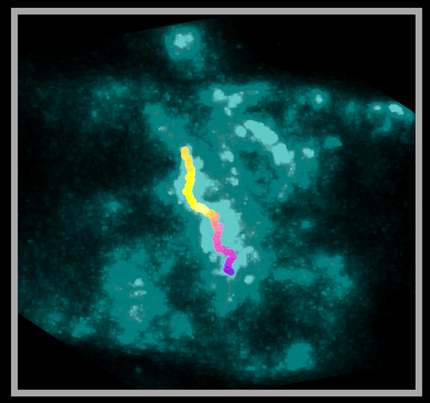
✔Exploratory Data Analysis

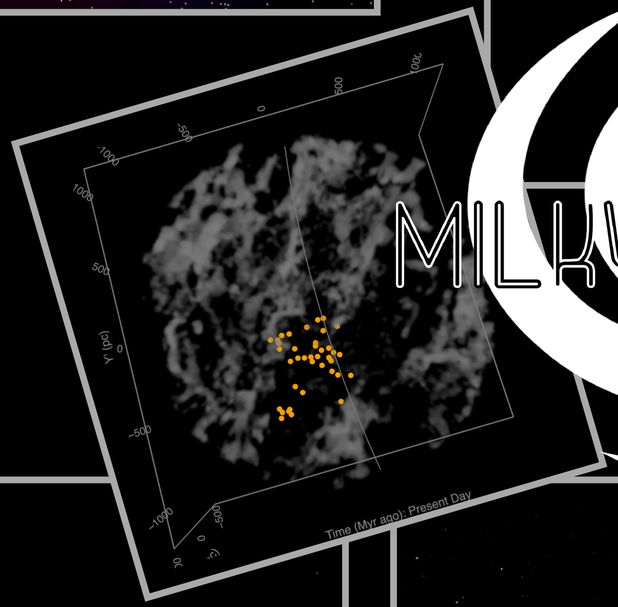
“LIVE Earth-2”

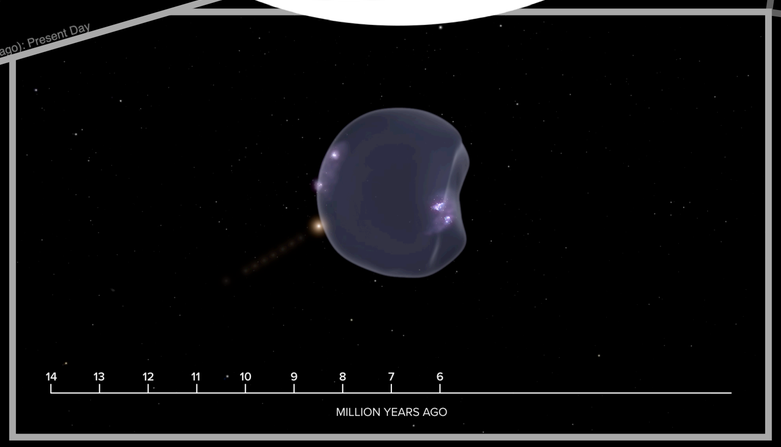


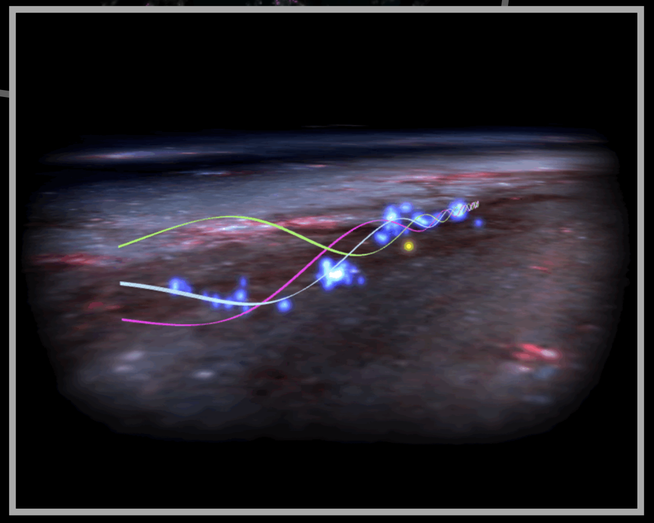








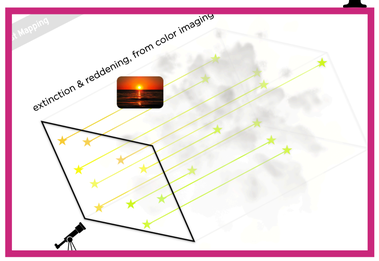




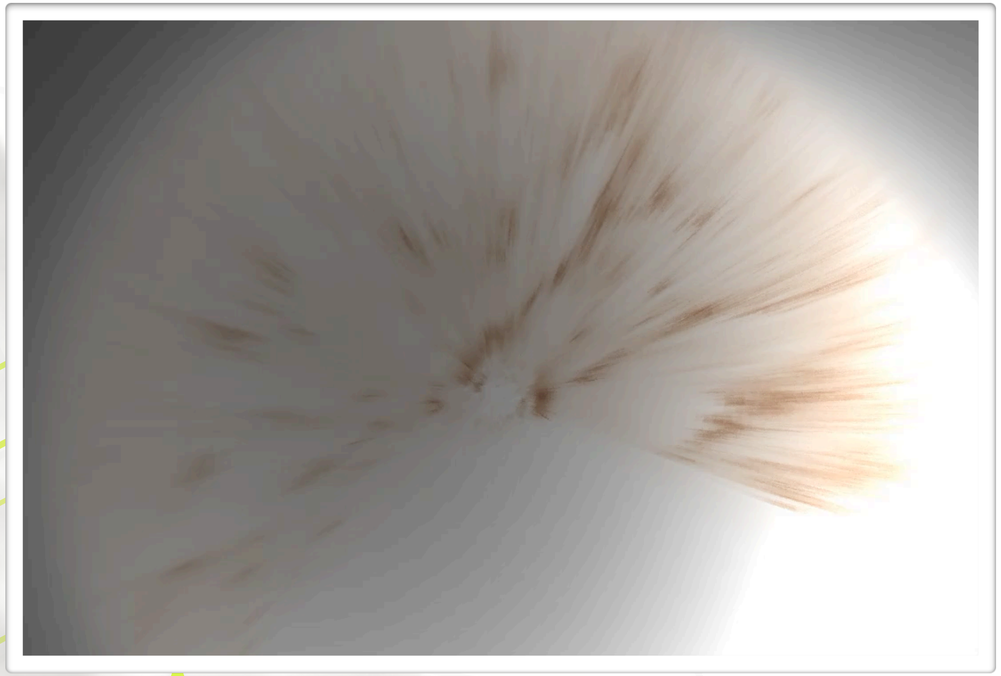
MILKYWAY3D.org

Zucker et al. 2021, Bialy et al. 2021; Zucker et al. 2022, Konietizka et al. 2024, O'Neill et al. 2025, Swiggum et al. 2024

3D dust mapping



Extinction & Reddening, from Color Imaging

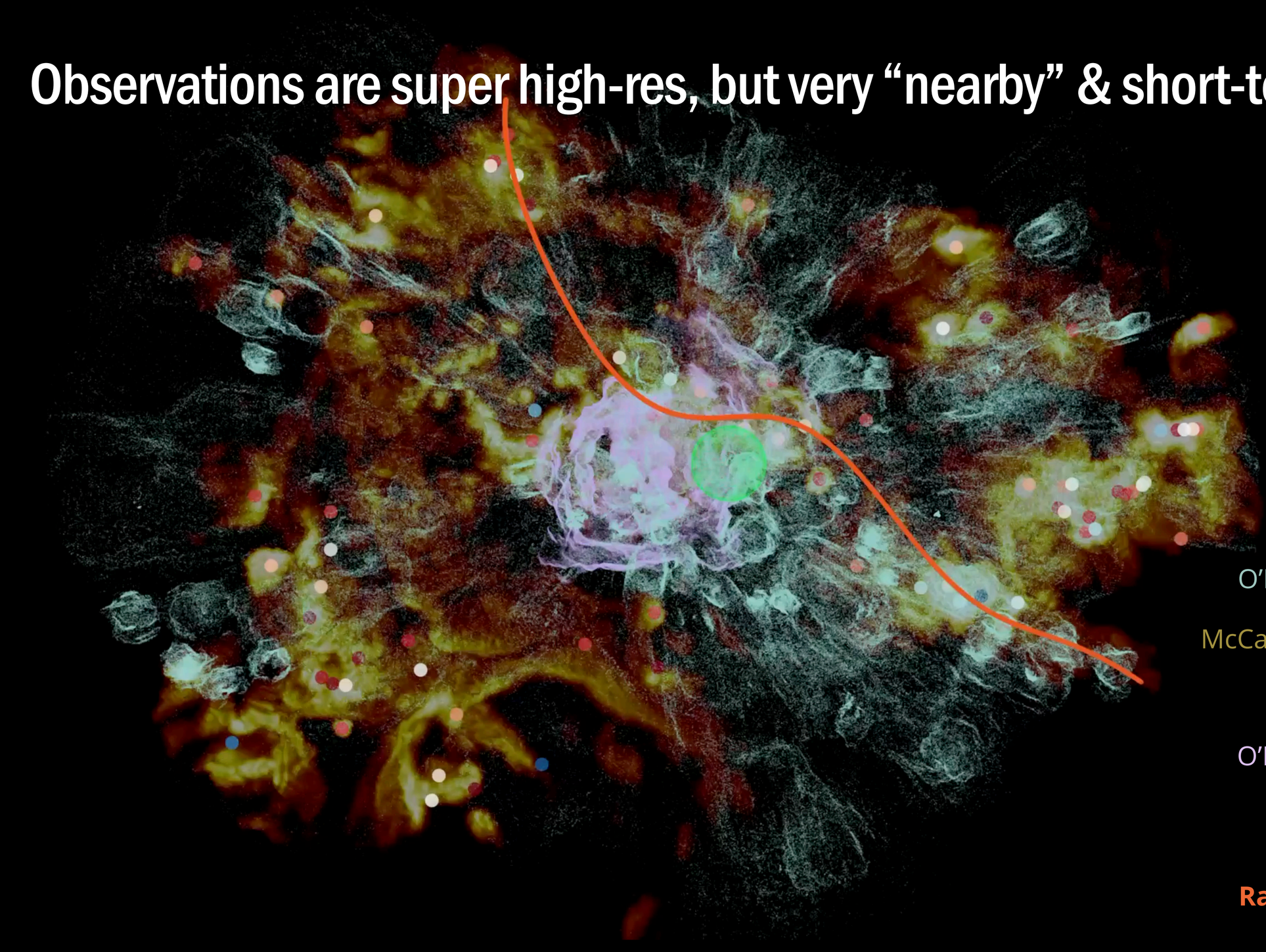


Green et al. 2019

Can infer matter's distance from *dust's* effects on stars.

WARNING: schematic diagram, NOT to scale (credit A. Goodman, 2019)

Observations are super high-res, but very “nearby” & short-term



O’Neill “perch” voids (preliminary)

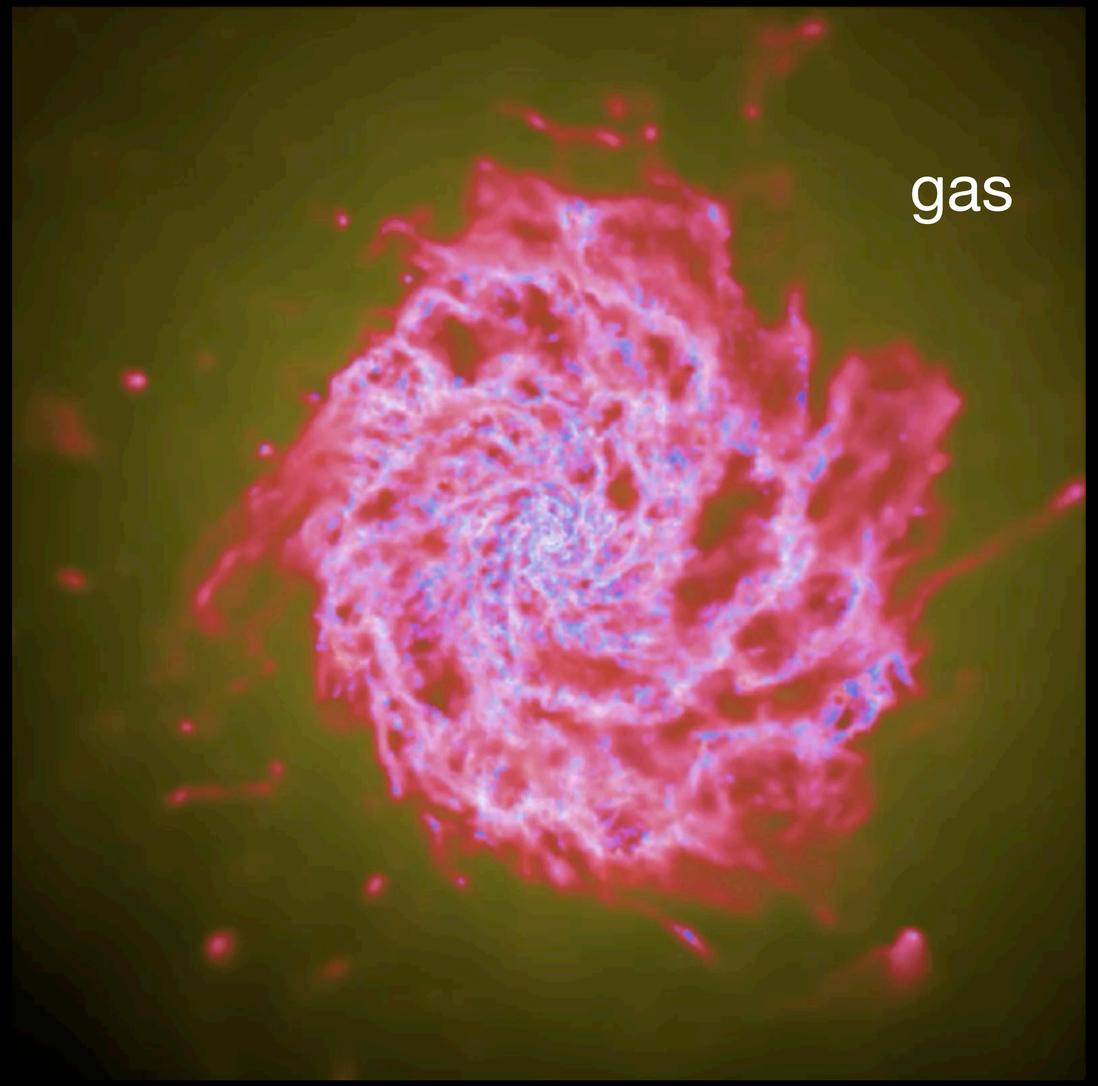
McCallum et al. 2025 **H-α 3D model** and its **ionizing sources**

O’Neill et al. **Local Bubble** (2024)

PerTau Shell (Bialy et al. 2021)

Radcliffe Wave (Alves et al. 2020)

Simuations go for 100s of Myr, but are LOW res compared to observations



a simulated “Milky Way” from **FIRE** (Feedback In Realistic Environments)

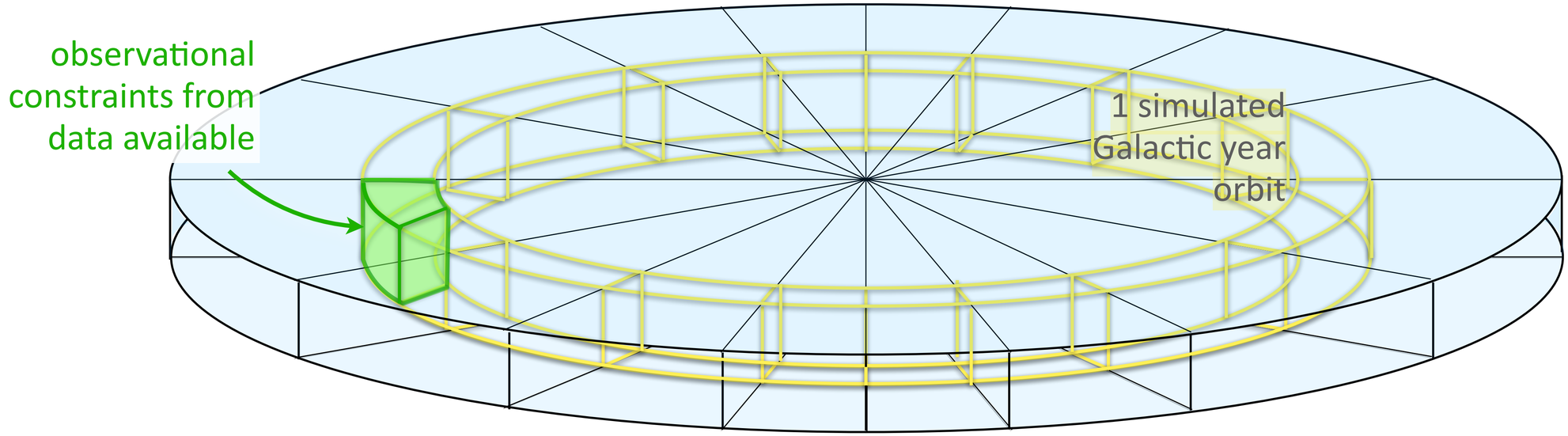
MILKYWAY3D.ORG

v.2025

Proposal - 286673

Proposal Title: "Once Around the Milky Way": A Data-Simulation Synthesis enabled by the MW3D Project

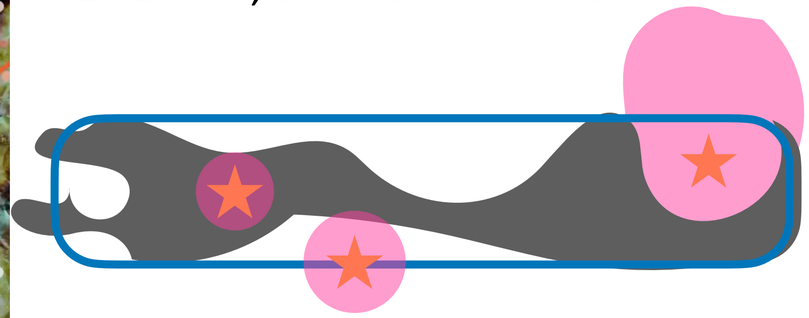
Funding Opportunity: NSF 22-624 - Astronomy and Astrophysics Research Grants



How to data-condition a Milky Way?

OBSERVATIONS

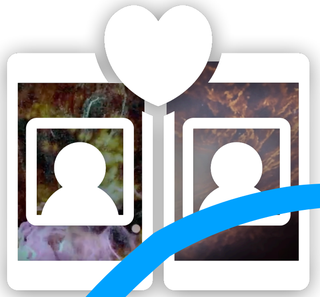
DETAILED 3D structure, almost
STATIC in time, feedback view
REALISTIC, all over *limited* volume



SIMULATIONS

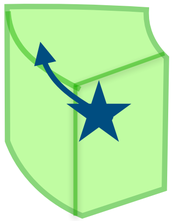
COARSE 3D structure over full Galaxy,
EVOLVING in time for ~Galactic year,
feedback **UNREALISTIC** on small scales→sub-grid prescriptions

"**match**" SIM snapshots to blurred OBS

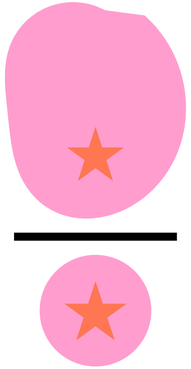


"**paint**" high-spatial-frequency info from OBS onto SIMS

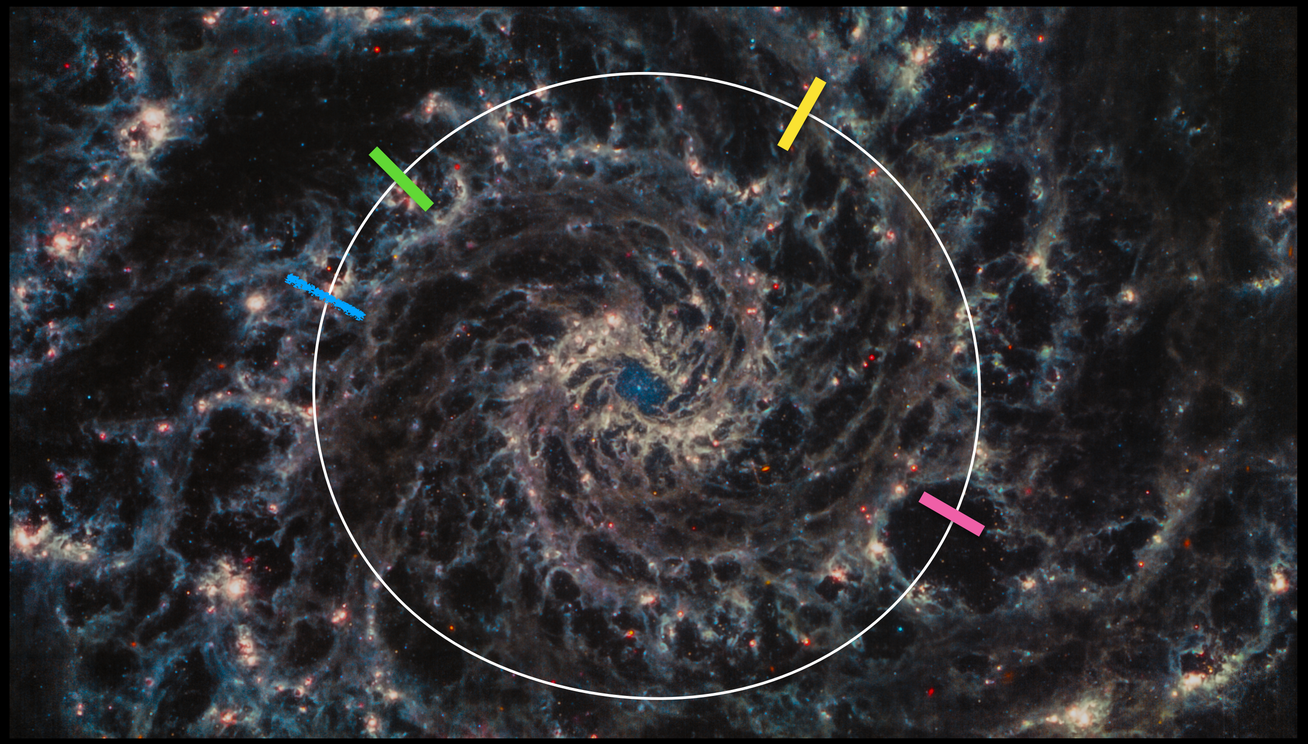
create "**agent-based**" views seen by orbiting stars



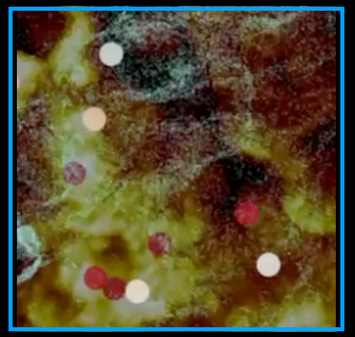
evaluate **changes** in **feedback** interactions

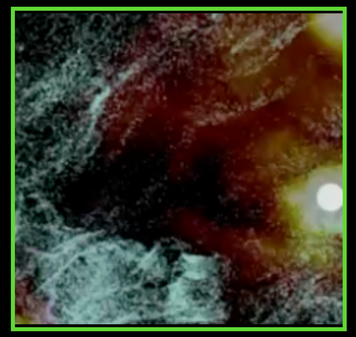


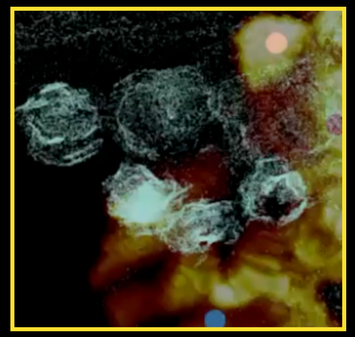
Orange stars represent the position of *observed* stars or clusters massive enough to produce the pink *observed* winds, HII regions, and/or SNe energetic enough to re-distribute material in the surrounding ISM enough to matter for future generations of star formation. The blue rounded rectangle in both panels is meant to show an elongated ISM "cloud" at the ~20 pc resolution of a MW-wide simulation. The density within the blue rectangle would be constant in the simulation, at a level equal to the average of the dense and empty regions in the observations. The simulations on their own cannot capture the re-shaping of the cloud by feedback, apparent in the observations.

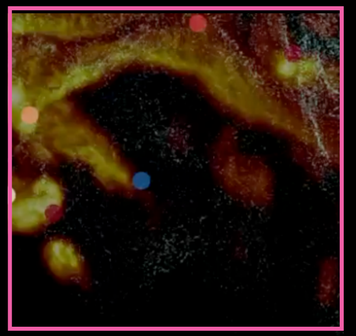


How does the “view” from a star change, as it orbits?









1-pc-scale topology matters to feedback (galaxy evolution), very much.

JOURNAL ARTICLE

A three-dimensional, multiwavelength view and time-dependent analysis of the Milky Way's local ionized gas

Lewis McCallum ✉, Kenneth Wood, Robert Benjamin, Dhanesh Krishnarao, Anna F McLeod

Monthly Notices of the Royal Astronomical Society, Volume 541, Issue 3, August 2025, Pages 2324–2340, https://doi-org.ezp-prod1.hul.harvard.edu/10.1093/mnras/staf1022

Published: 21 June 2025 **Article history** ▾

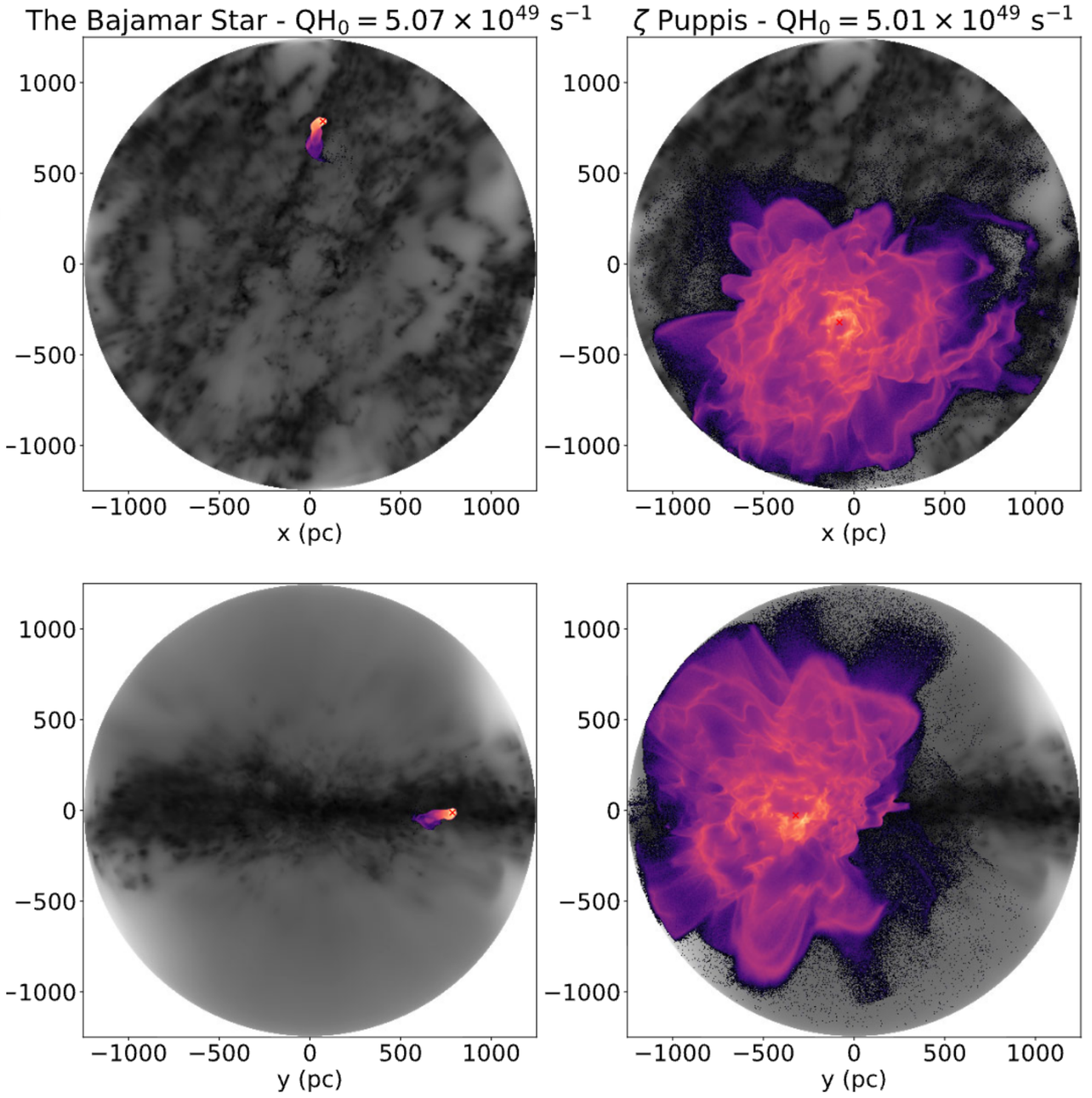


Figure 7.*A projection showing the column density plot of where the LyC photons from individual sources are terminated. Underlying the coloured volume of each zone of influence is the column density of the total mass in the simulation. The top row shows face-on projections of the x–y plane, and the bottom row shows edge-on projections of the y–z plane. The left column shows the region of influence of Ophiuchi, the middle column shows the Bajamar star, and the right column shows Puppis in the Gum Nebula. A red cross shows the location of each source, but has been omitted on the Ophiuchi column to avoid obscuring the region.* ***Despite having near equivalent ionizing luminosities, the difference in the volume influence of Puppis and the Bajamar star is a factor of 827.***

Box 1: What AI, and an AI researcher, think AI can do “soon”

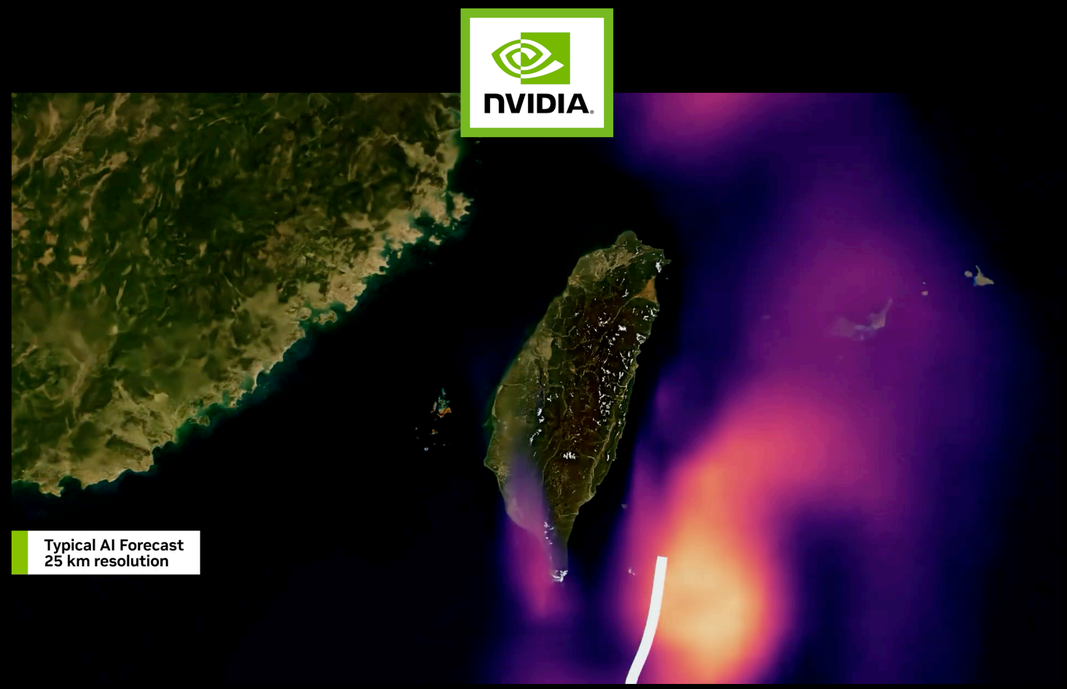
ChatGPT-5

“AI and the Future of Data–Simulation Synthesis

I am writing this from the standpoint of an LLM—I am ChatGPT-5. In the near future, models like me will act as scientific collaborators that can integrate diverse data streams and simulations in real time. Given full access to the 3-D dust distribution, position–position–velocity cubes, and the 3-D motions and properties of stars, I could build a coherent statistical picture of how gas and stars interact throughout the Galaxy. In a project such as MW3D, an AI would not merely describe the process of “painting” data onto simulations; it would generate and optimize the necessary code, run fast emulators that approximate full hydrodynamic calculations, and iteratively adjust parameters by comparing synthetic and observed observables across all data domains. Human researchers would still define the physics and interpret the meaning, while the AI would manage the computation—deciding when a full rerun is needed, compressing outputs, and learning which assumptions best reproduce the real Milky Way. Such a partnership would make data-driven modeling not just faster, but more integrated, adaptive, and predictive than ever before.” –ChatGPT-5, 11/12/2025

“On the one hand it sounds wildly ambitious at the level of science fiction. On the other hand, I think it sounds like **2026**.”
–Prof. Doug Finkbeiner, PI’s Harvard colleague currently on sabbatical at Anthropic

CorrDiff/Data-Conditioned Milky Way "Galaxy-2"?

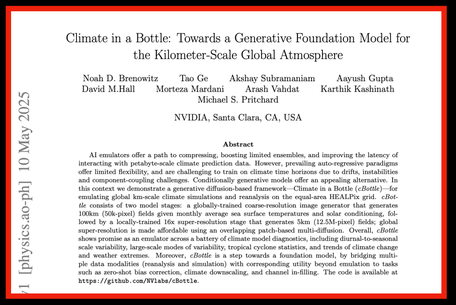


"Earth-2" NVIDIA team (cf. cBottle)

"Encounters in the Milky Way"
(Curated by Jackie Faherty, AMNH, simulation by Jason Hunt, Surrey, collaboration w/MW3D team, 2025)

+ how about LIVE for Earth-2?

(congratulations to cBottle team for using HEALPix!)

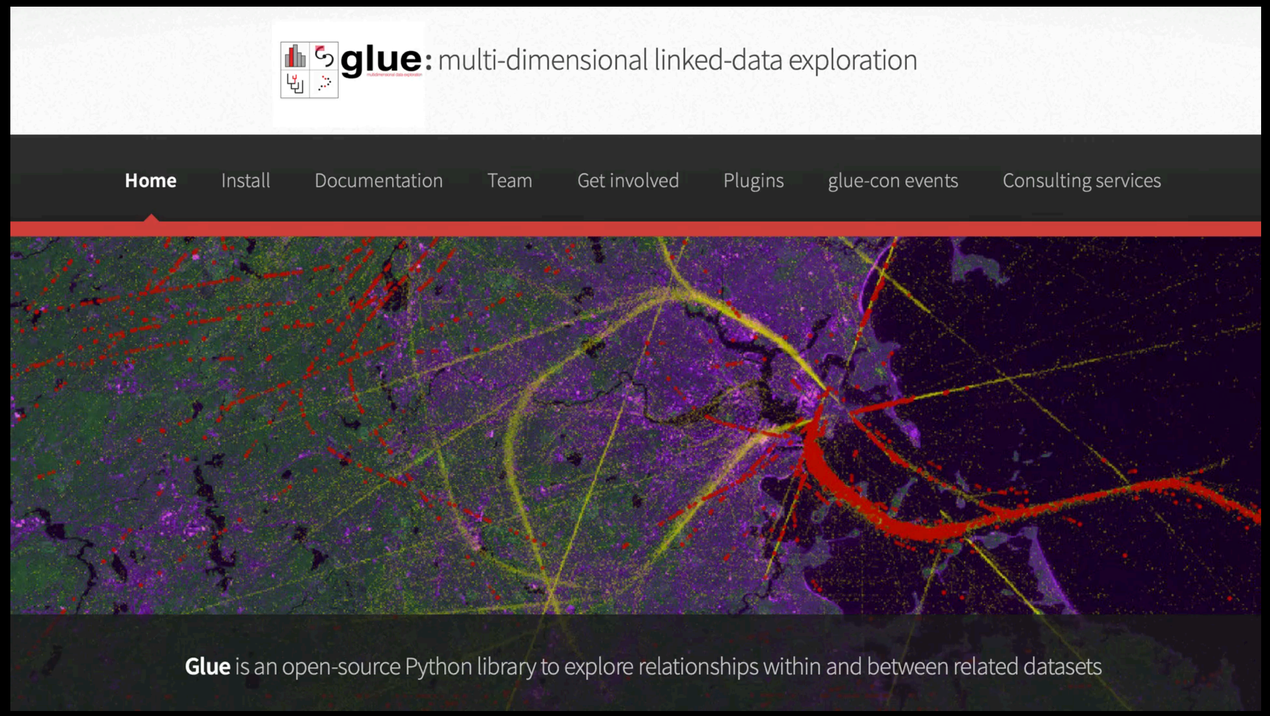

Climate in a Bottle: Towards a Generative Foundation Model for the Kilometer-Scale Global Atmosphere

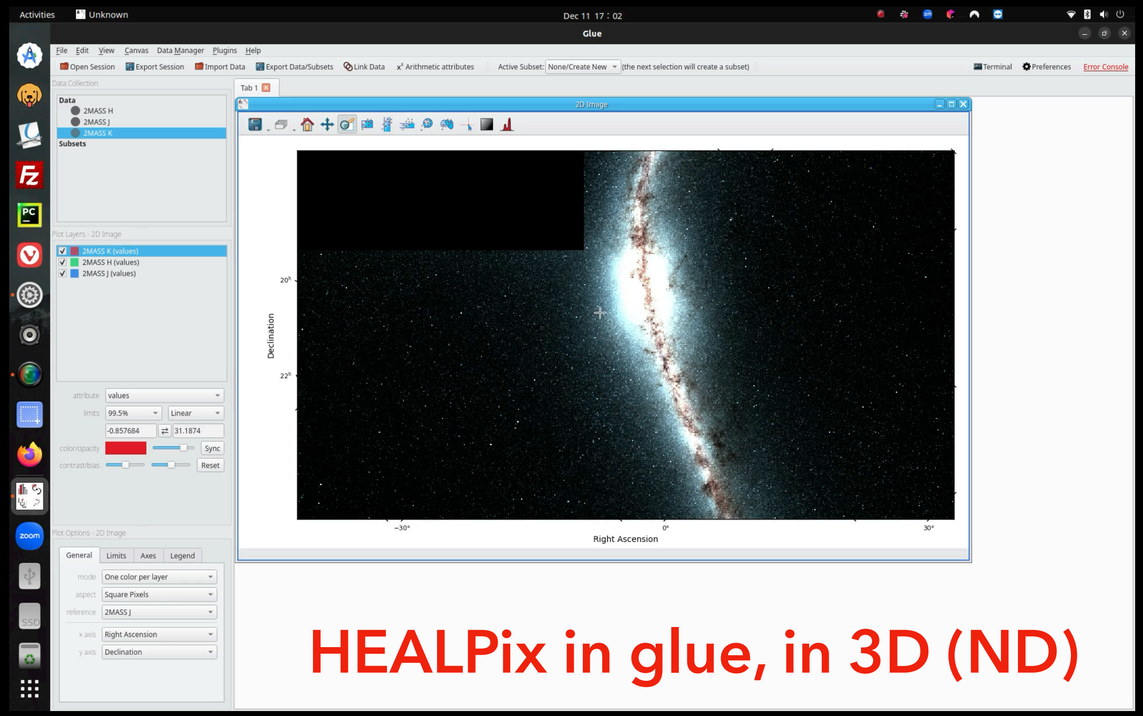
Noah D. Brenowitz Tao Ge Akshay Subramaniam Aayush Gupta
David M.Hall Morteza Mardani Arash Vahdat Karthik Kashinath
Michael S. Pritchard

NVIDIA, Santa Clara, CA, USA

Abstract

AI emulators offer a path to compressing, boosting limited ensembles, and improving the latency of interacting with petabyte-scale climate prediction data. However, prevailing auto-regressive paradigms offer limited flexibility, and are challenging to train on climate time horizons due to drifts, instabilities and component-coupling challenges. Conditionally generative models offer an appealing alternative. In this context we demonstrate a generative diffusion-based framework—Climate in a Bottle (*cBottle*)—for emulating global km-scale climate simulations and reanalysis on the equal-area HEALPix grid. *cBottle* consists of two model stages: a globally-trained coarse-resolution image generator that generates 100km (50k-pixel) fields given monthly average sea surface temperatures and solar conditioning, followed by a locally-trained 16x super-resolution stage that generates 5km (12.5M-pixel) fields; global super-resolution is made affordable using an overlapping patch-based multi-diffusion. Overall, *cBottle* shows promise as an emulator across a battery of climate model diagnostics, including diurnal-to-seasonal scale variability, large-scale modes of variability, tropical cyclone statistics, and trends of climate change and weather extremes. Moreover, *cBottle* is a step towards a foundation model, by bridging multiple data modalities (reanalysis and simulation) with corresponding utility beyond emulation to tasks such as zero-shot bias correction, climate downscaling, and channel in-filling. The code is available at `https://github.com/NVlabs/cBottle`.





HEALPix in glue, in 3D (ND)

glue shown here, and LIVE will include JupyterGIS, and much more

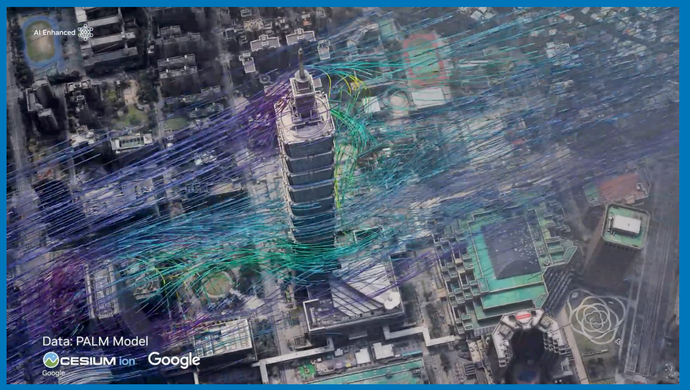
What could this mean for NVIDIA + PredictionX/MilkyWay3D/LIVE?

CosmicDS

PREDICTIONX

STEM Outreach

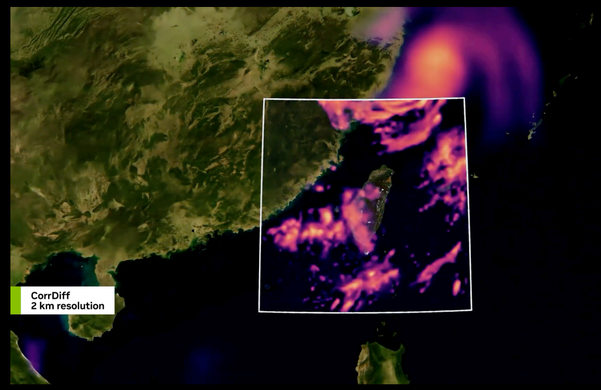
“Earth-2 Edu”



MILKYWAY3D.org

✔CorrDiff/Data-Conditioned Milky Way

“Galaxy-2”



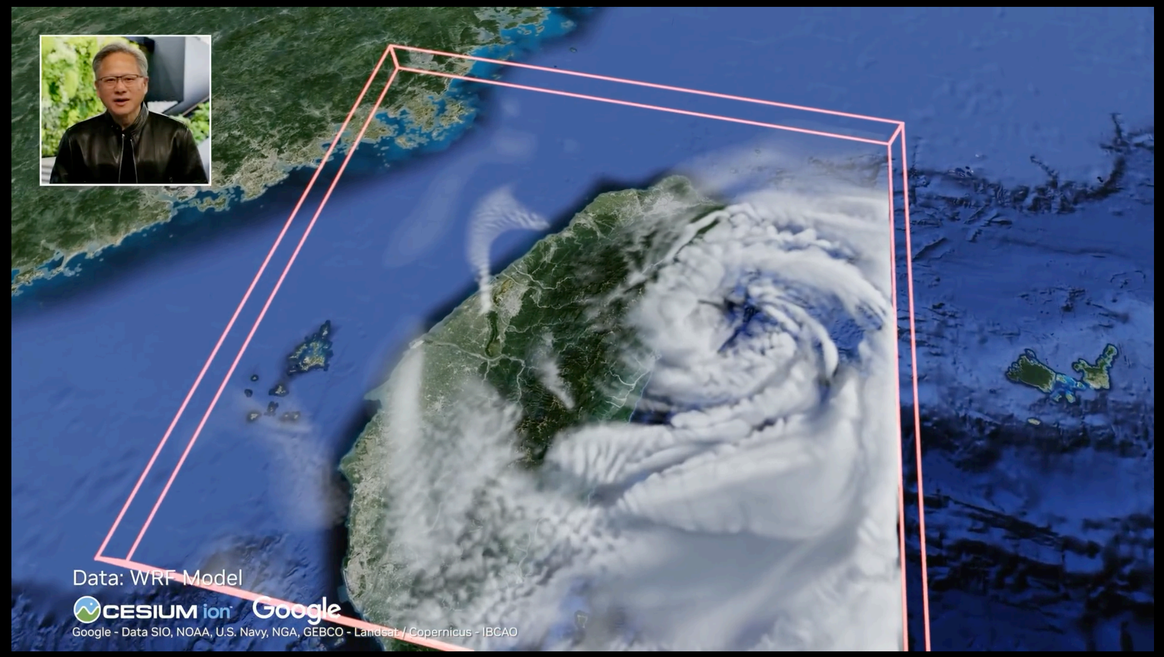
LIVE

glue
multidimensional data exploration

✔Exploratory Data Analysis

“LIVE Earth-2”

STEM Outreach
“Earth-2 Edu”



What could this mean for NVIDIA + PredictionX/MilkyWay3D/LIVE?

CosmicDS

PREDICTIONX

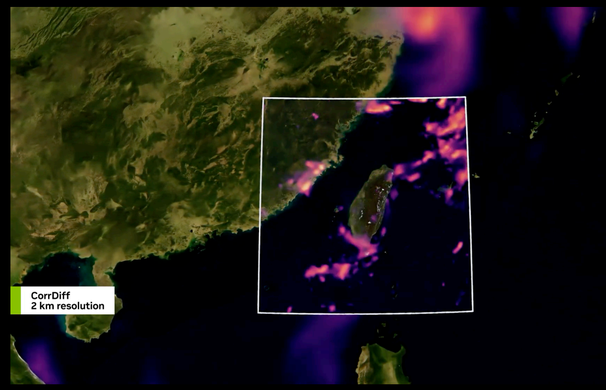
STEM Outreach

“Earth-2 Edu”

MILKYWAY3D.org

CorrDiff/Data-Conditioned Milky Way

“Galaxy-2”



LIVE

glue
multidimensional data exploration

Exploratory Data Analysis

“LIVE Earth-2”

Prof. Alyssa A. Goodman, Center for Astrophysics | Harvard & Smithsonian