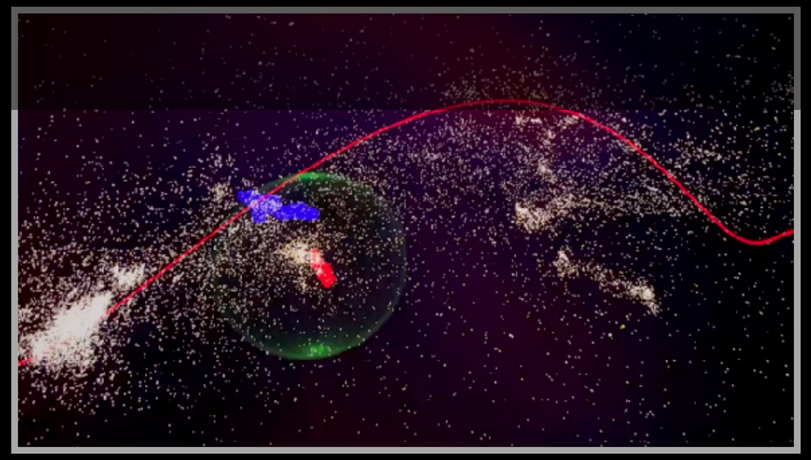
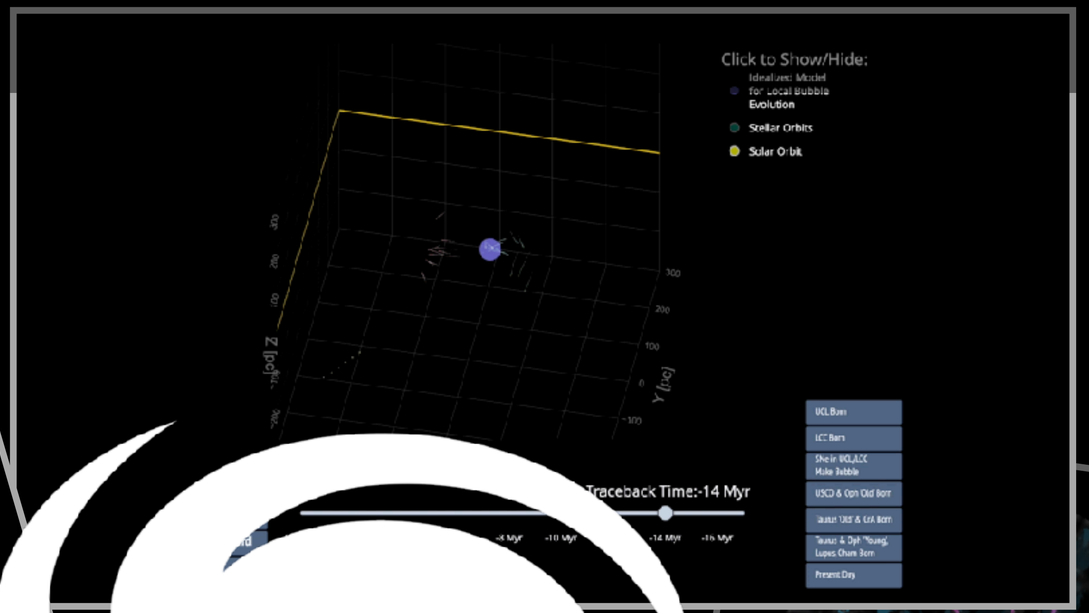
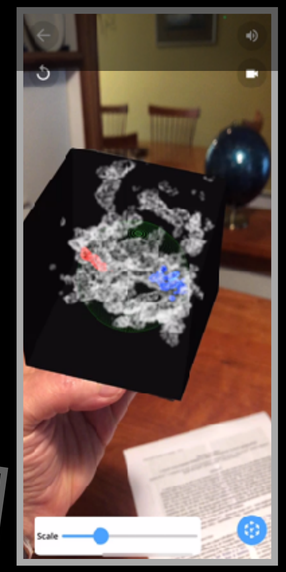
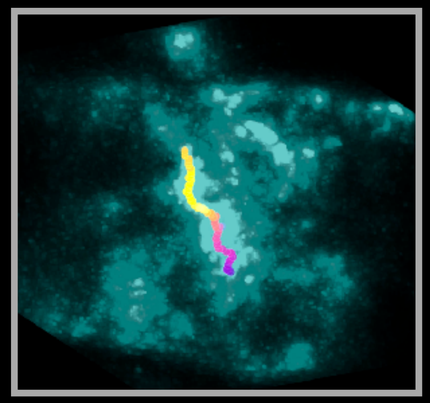
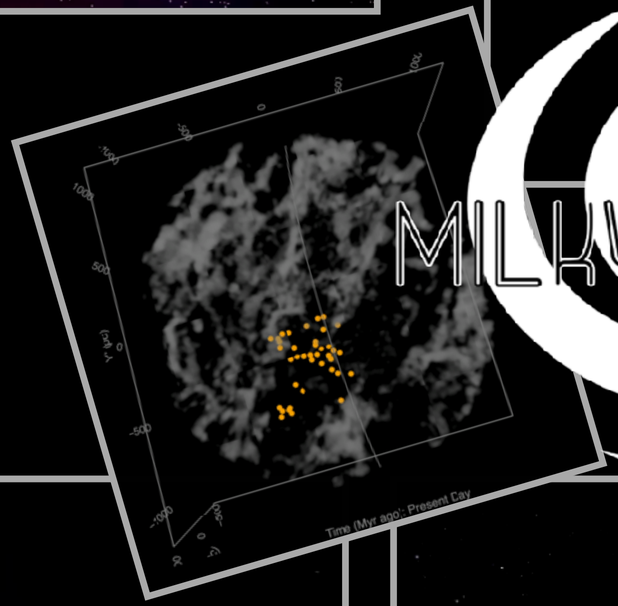
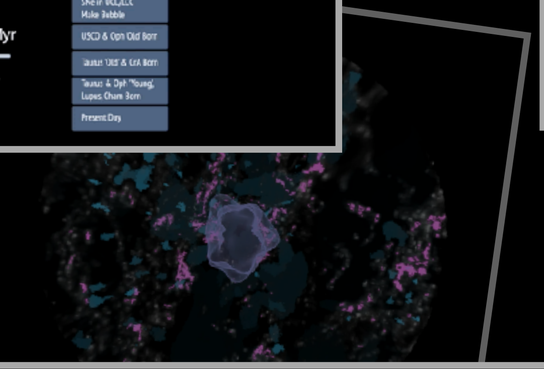




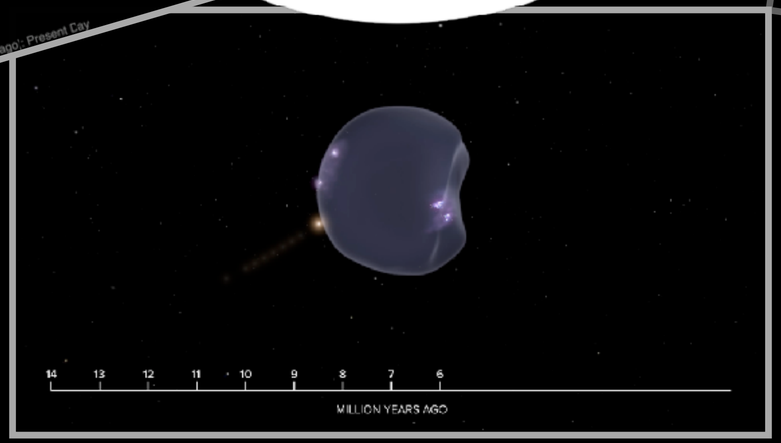


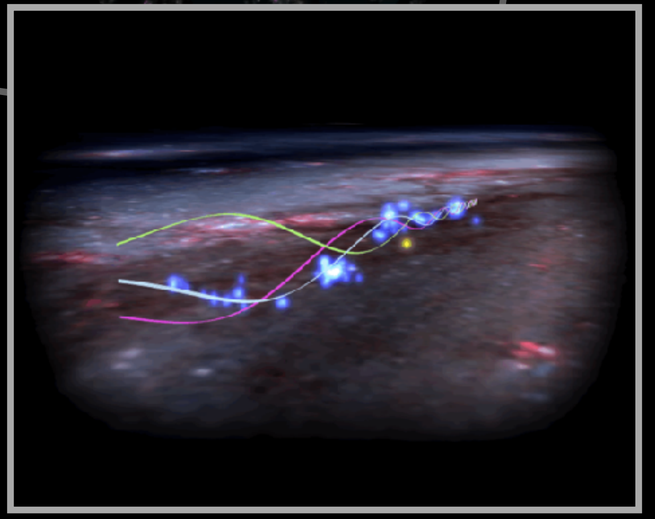






MILKYWAY3D.org



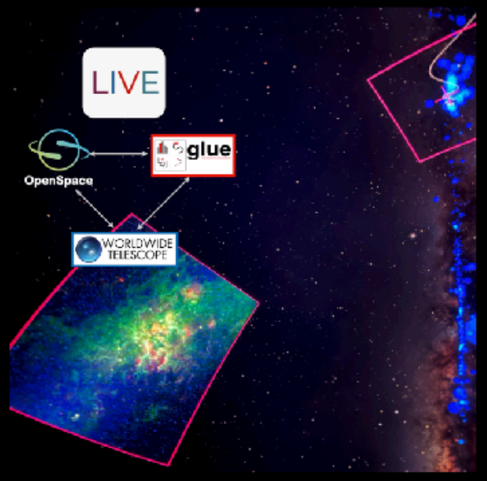


Zucker et al. 2021, Bialy et al. 2021; Zucker et al. 2022, Konietizka et al. 2024, O'Neill et al. 2025, Swiggum et al. 2024

Skip directly to: DATA | SOFTWARE | OUTREACH | JOINING IN

Science

Researchers worldwide contribute to, and use, MilkyWay3D resources in order to improve our understanding of the Milky Way's origins, workings, and future.



Data & Software

We combine data and software that can offer myriad views of the Milky Way, as best we know it. The software is built using the LIVE Environments exploratory data analysis tools.

MW3D for Everyone

Using the same data and tools as researchers, educators and learners of all ages can explore and study the Galaxy in which we all live.

MilkyWay3D's open science premise: Data + Software + People + Computing + Physics = Understanding, Learning, and Discovery

MILKYWAY3D.org

Bringing you the Galaxy as we know it. Updated as we learn more.

MILKYWAY3D.ORG

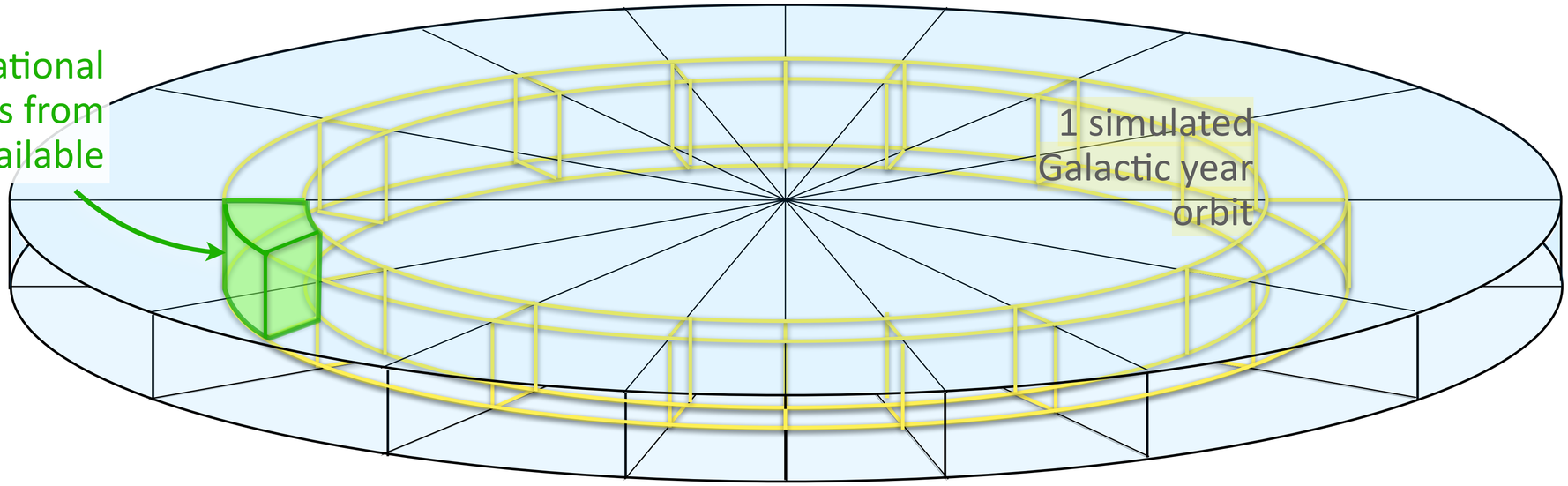
v.2025

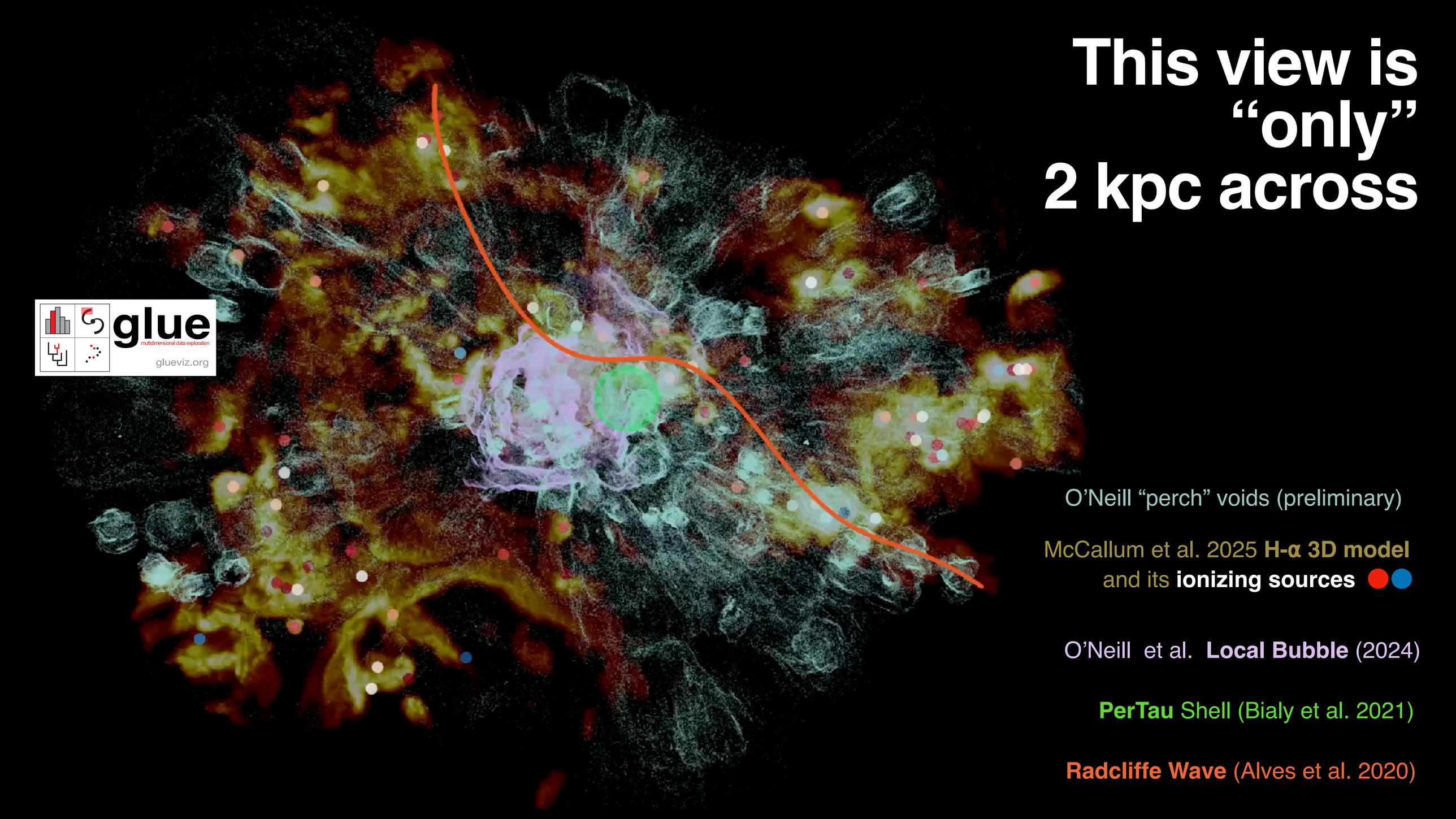
Proposal - 286673

Proposal Title: "Once Around the Milky Way": A Data-Simulat oject

Funding Opportunity: NSF 22-624 - Astronomy and Astrophysics

Please ask me about this!





This view is “only” 2 kpc across

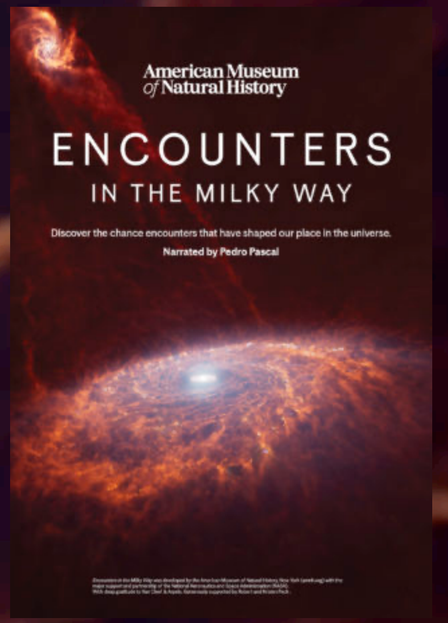
O’Neill “perch” voids (preliminary)

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

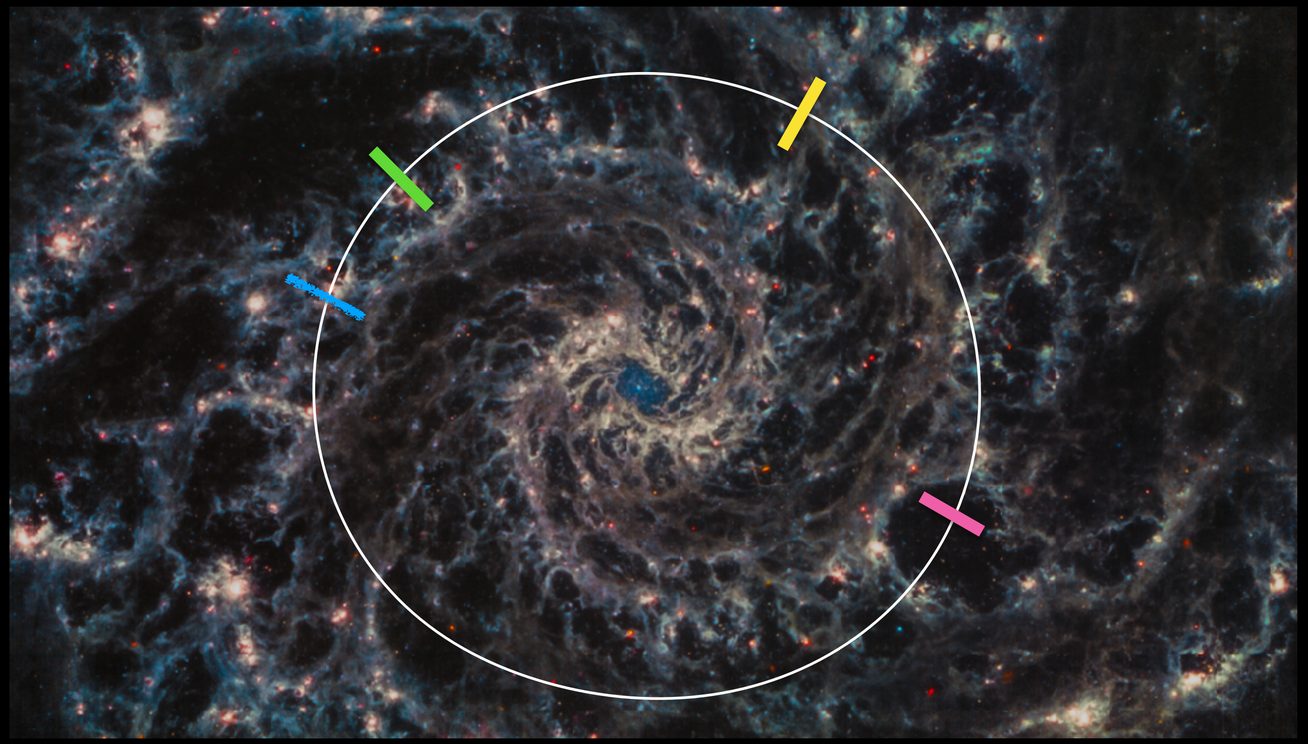
O’Neill et al. **Local Bubble** (2024)

PerTau Shell (Bialy et al. 2021)

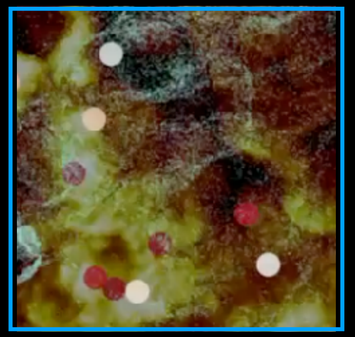
Radcliffe Wave (Alves et al. 2020)

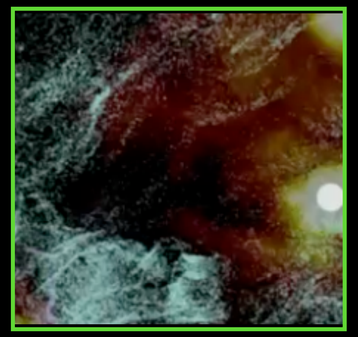


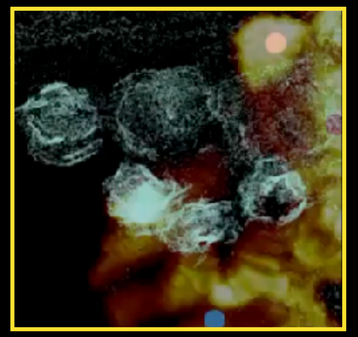
rator: Jackie Faherty, AMNH, Simulation Jason Hunt, Surrey, plus MW3D collaborators

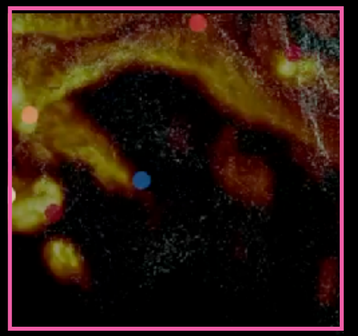


How does the “view” from a star change, as it orbits?









Yes, 1-pc-scale topology matters to feedback, 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** ▾

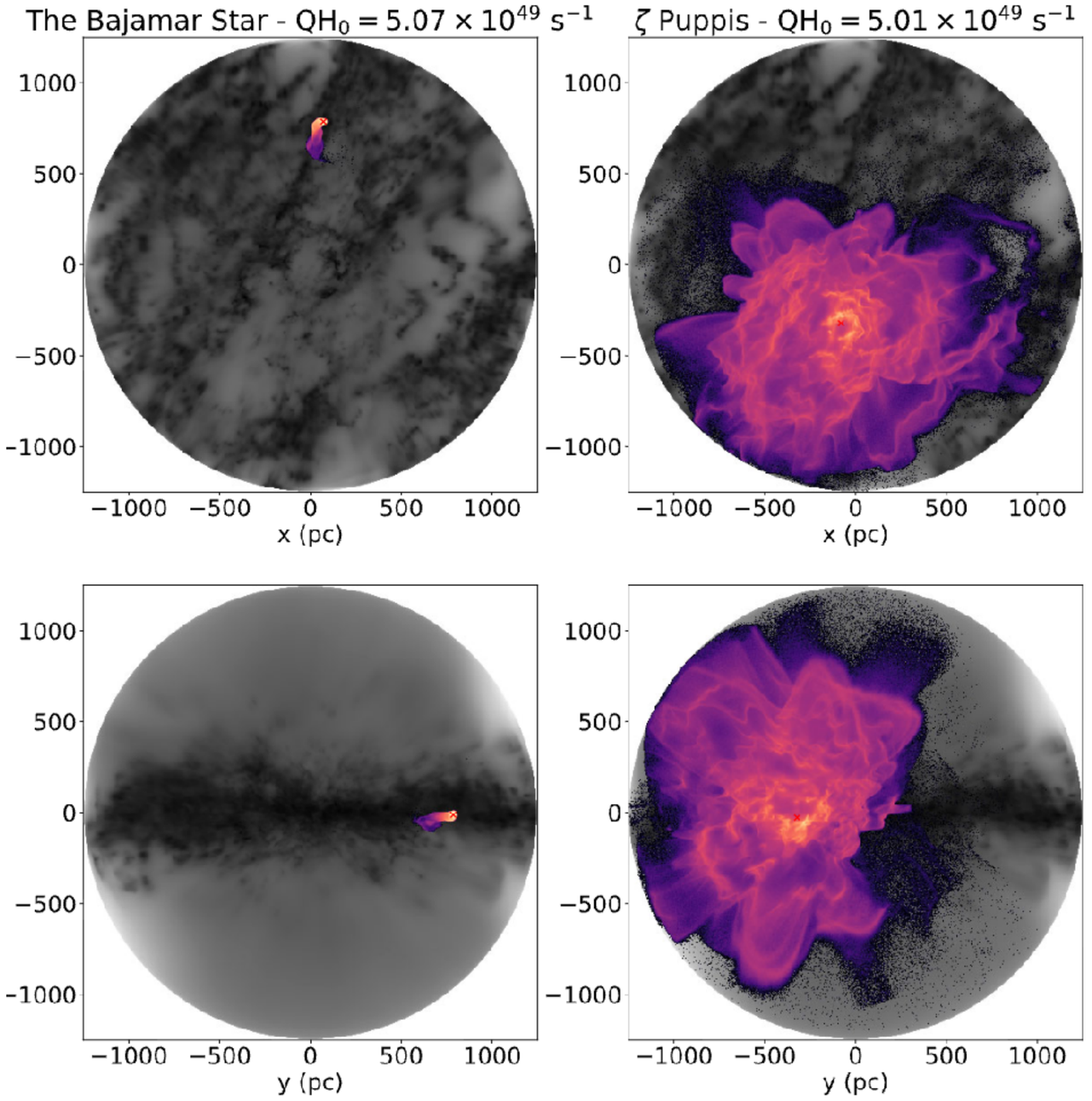
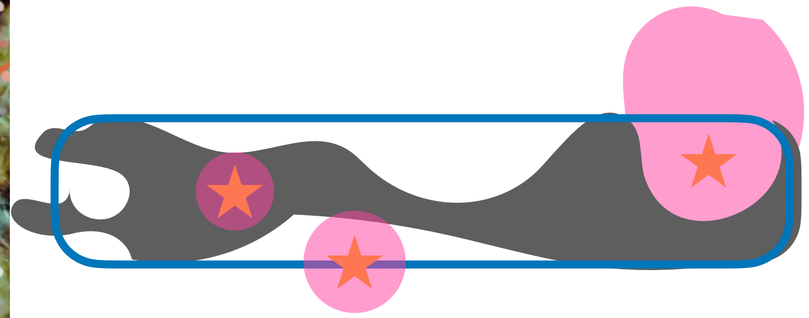


Figure 7.*A projection showing the column density plot of where the LyC photons from individual sources are terminated. Underlying the coloured volume of each zone of influence is the column density of the total mass in the simulation. The top row shows face-on projections of the x–y plane, and the bottom 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.***

The whole “plan”

OBSERVATIONS

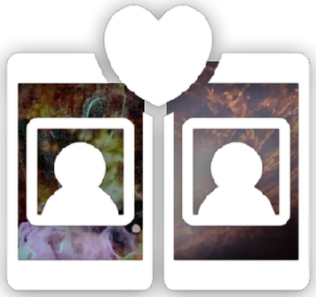
DETAILED 3D structure, almost
STATIC in time, feedback view
REALISTIC, all over *limited* volume

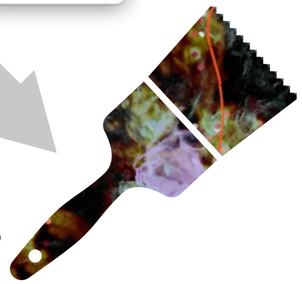


SIMULATIONS

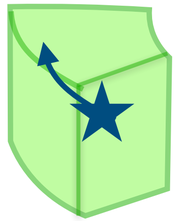
COARSE 3D structure over full Galaxy,
EVOLVING in time for ~Galactic year,
feedback **UNREALISTIC** on small
scales→sub-grid prescriptions

“**match**” SIM snapshots to blurred OBS

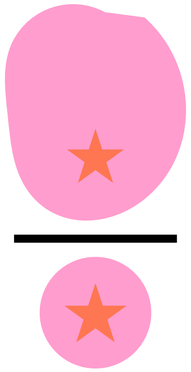




“**paint**” high-spatial-frequency info from OBS onto SIMS

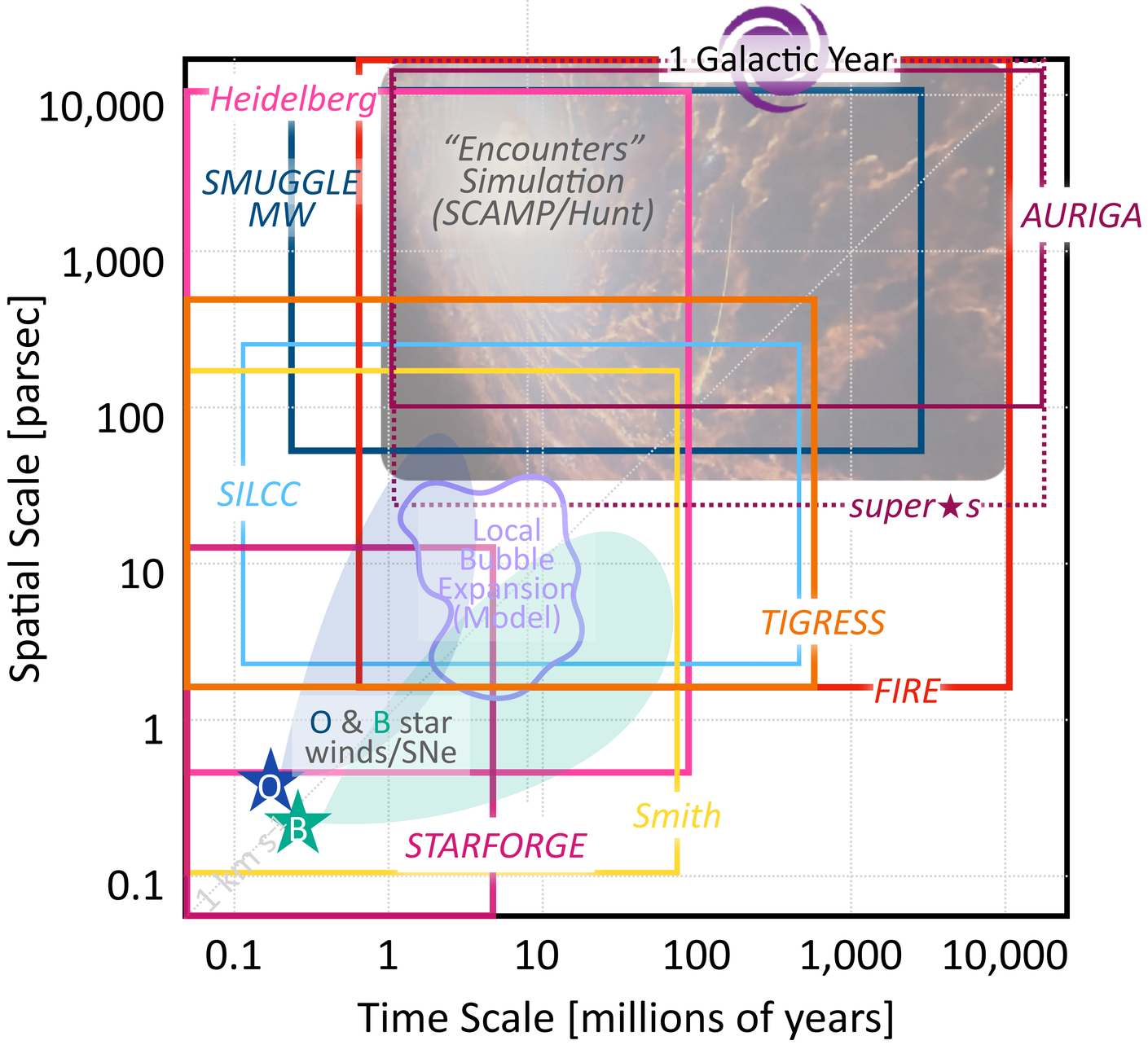


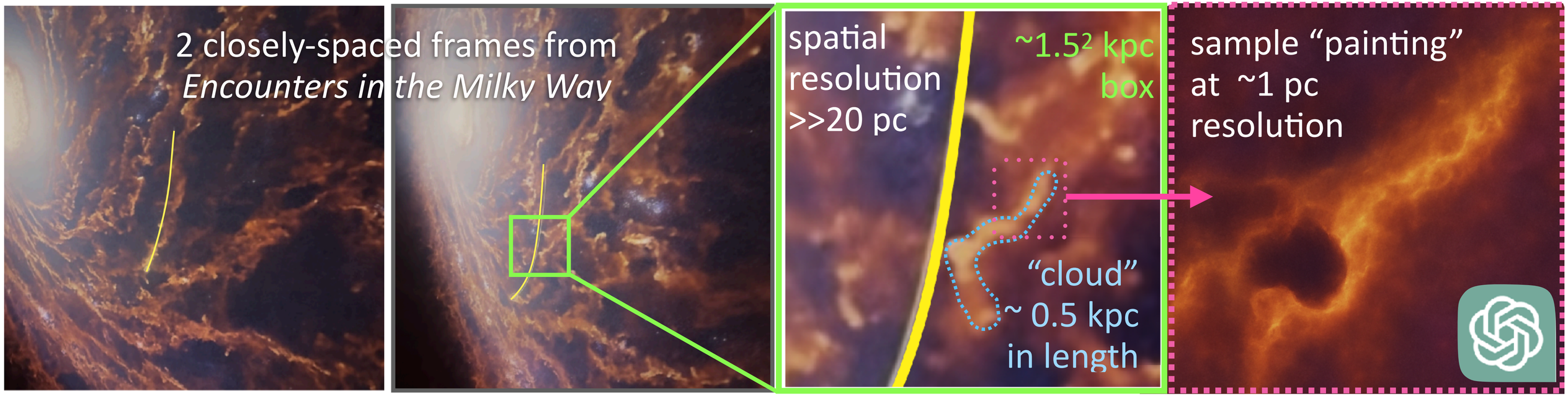
create “**agent-based**” views seen by orbiting stars



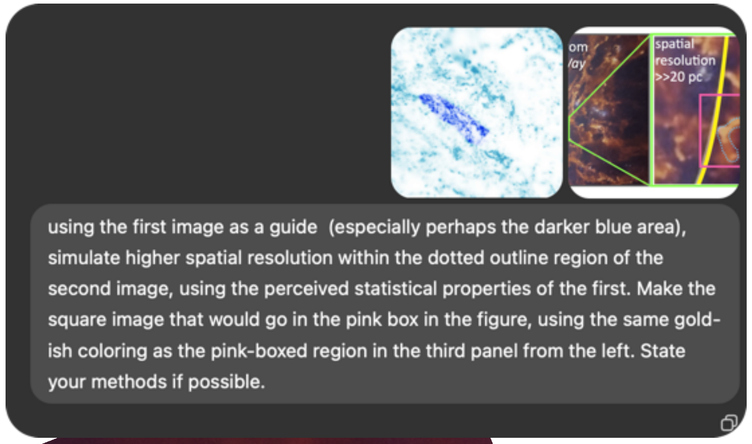
evaluate **changes** in **feedback** interactions

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.


1 Galactic Year
Heidelberg
SMUGGLE MW
"Encounters" Simulation (SCAMP/Hunt)
AURIGA
SILCC
super★s
Local Bubble Expansion (Model)
TIGRESS
FIRE
O & B star winds/SNe
O
B
Smith
STARFORGE
1 km s⁻¹
10,000
1,000
100
10
1
0.1
Spatial Scale [parsec]
0.1
1
10
100
1,000
10,000
Time Scale [millions of years]


2 closely-spaced frames from
Encounters in the Milky Way
spatial resolution >>20 pc
~1.5² kpc box
"cloud" ~ 0.5 kpc in length
sample "painting" at ~1 pc resolution

NY failures…with some success.



ChatGPT's code-based (concise "recipe")

Estimate the small-scale texture statistics from the blue reference image

Convert to grayscale, crop centrally to avoid edges.

Compute the isotropic radial power spectrum (FFT → $|F|^2$ → radial average).

This captures the "how much structure at each spatial frequency" of the high-res data.

Synthesize a new Gaussian random field with the same spectrum

Create a random complex field with the measured magnitude (√PSD as a function of radius) and random phases.

Inverse FFT → a new field whose spatial frequency content matches the reference.

Normalize and lightly band-pass filter to emphasize filamentary scales (e.g., subtract a σ≈5–6 px blur to bring out "wisps").

Imprint a long, squiggly overdensity to mimic the dotted "cloud"

Lay down a spline-like path through the image (length/curvature like the outline).

Add an anisotropic Gaussian tube along that path with a slowly varying width (to resemble clumpy substructure).

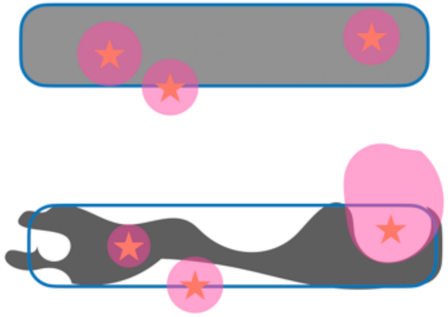
Modulate that tube by the high-frequency component of the synthesized field so it inherits realistic speckle/porosity.

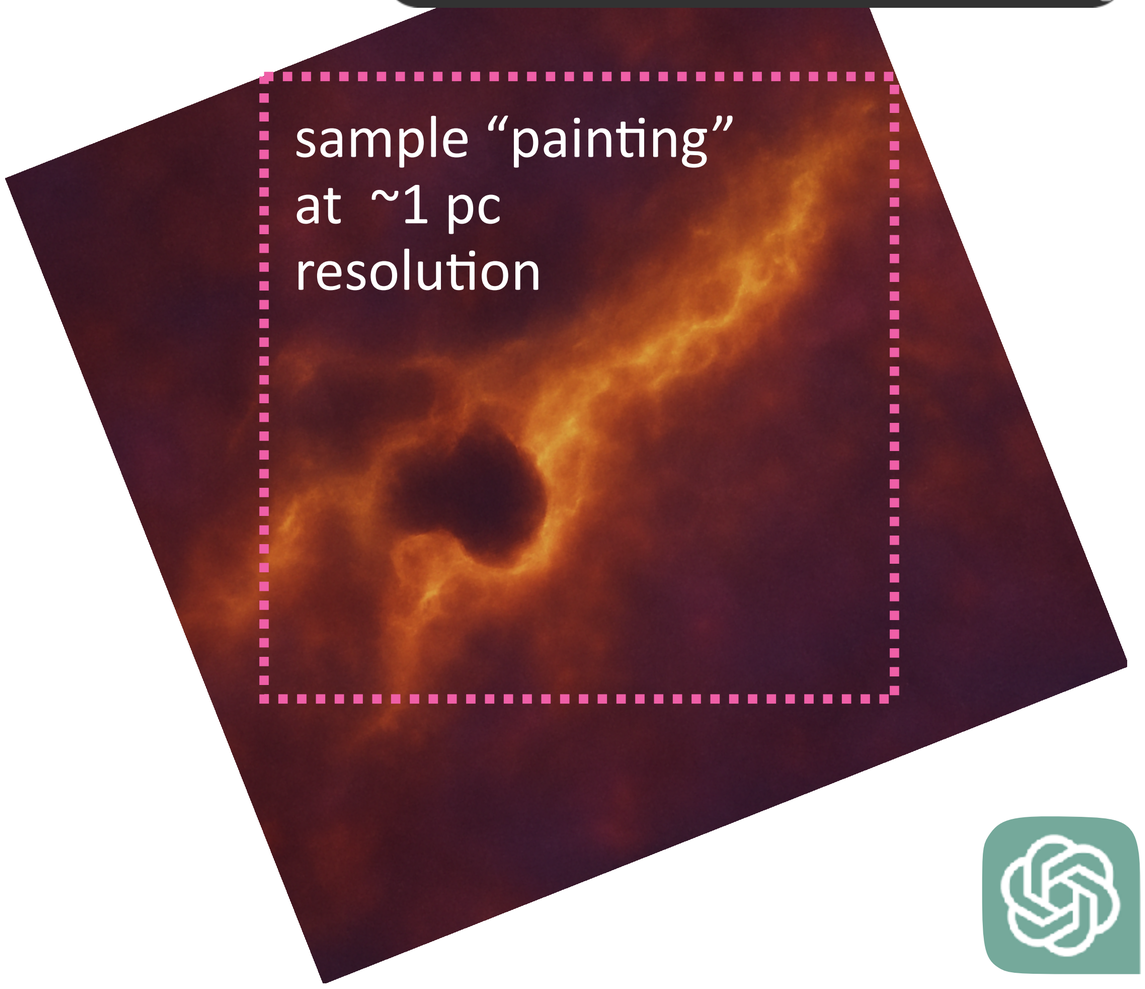
Contrast + color transfer to the gold palette

Local contrast stretch (clip low/high percentiles, e.g., 25th–99.5th) to keep punch without blowing highlights.

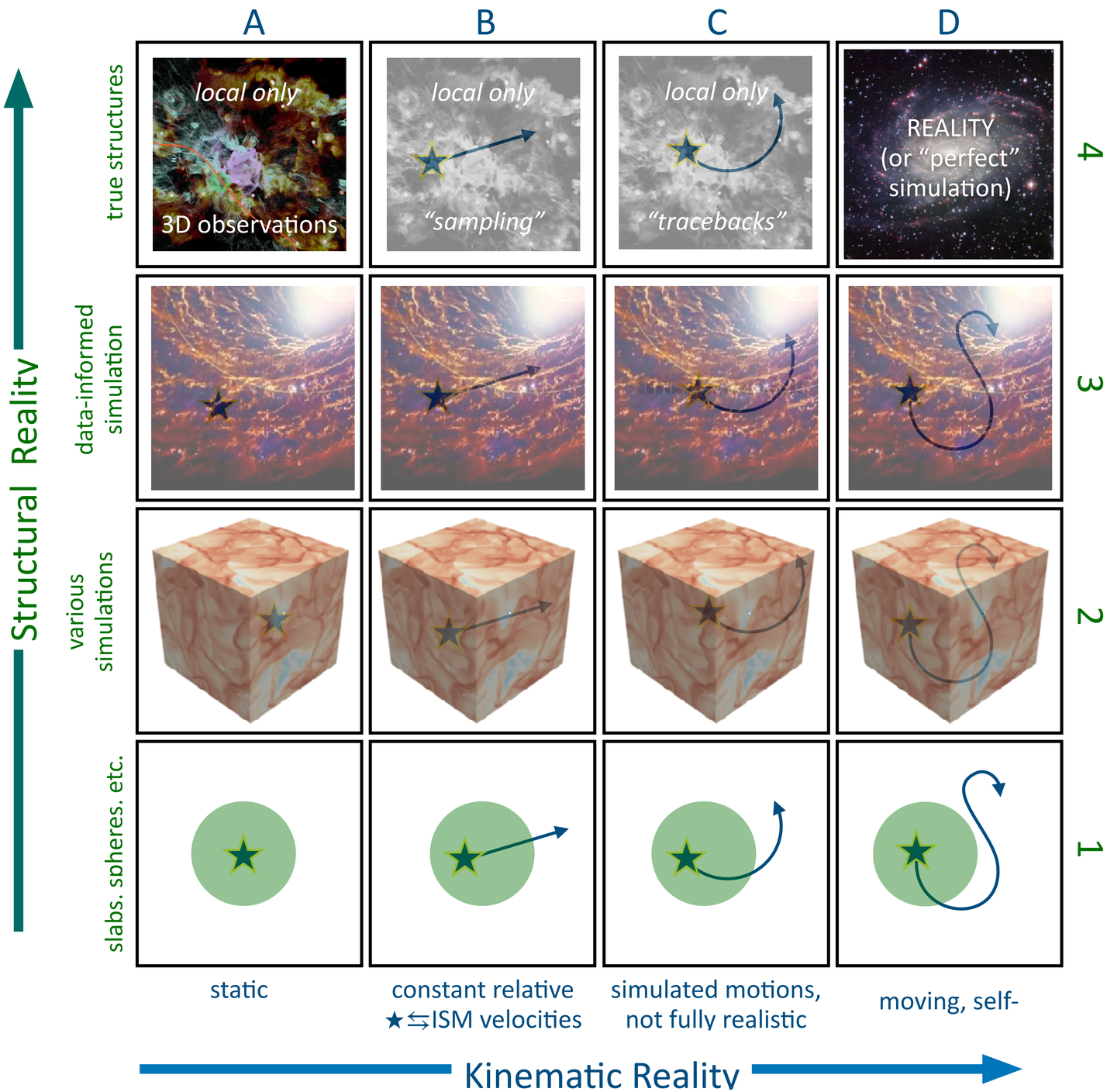
Apply a custom gold colormap (dark charcoal → warm brown → pale gold), or do LAB-space histogram matching to the gold patch from the third panel so the hue looks consistent.

Export a square PNG sized to the pink box.





How close to reality can we get?



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

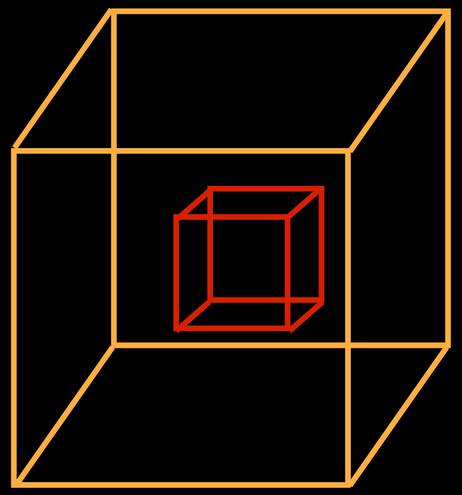
Exploratory Data Analysis...

What is it, and why does it matter in the age of big surveys and AI?

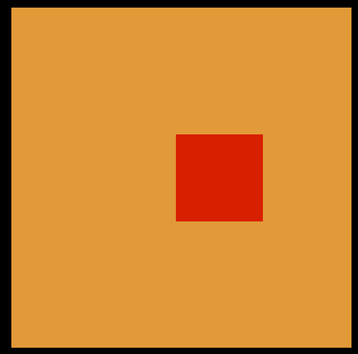
Alyssa A. Goodman
Center for Astrophysics | Harvard & Smithsonian

Exploratory Data Analysis (EDA)

"Linked Views of High-dimensional Data"

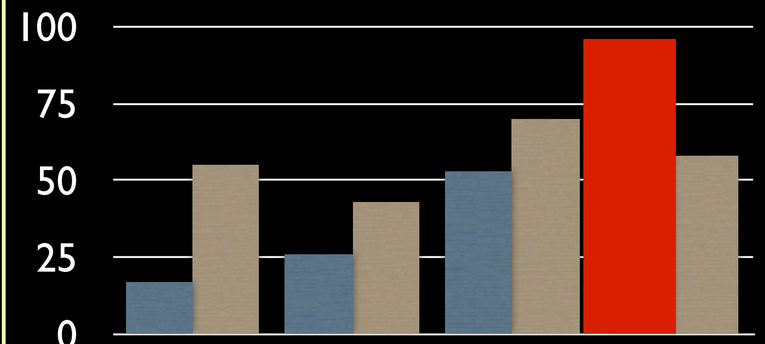


3D

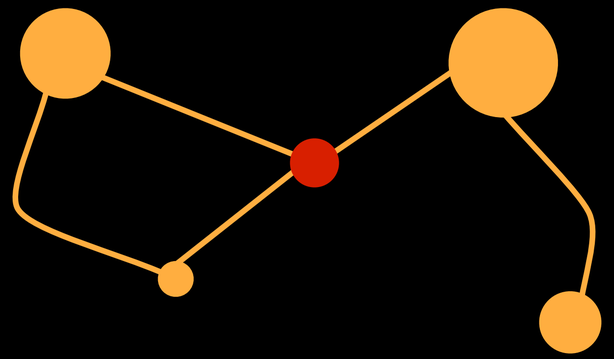


2D

Statistics



Data Abstraction



John Tukey
1973

glue
multidimensional data exploration

"today"

figure, by M. Borkin, reproduced from Goodman 2012, "Principles of High-Dimensional Data Visualization in Astronomy"

glue
multidimensional data exploration
glueviz.org

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)

multi-layer images, with additional points, curves, and selection layers

selections across all views

standard x-y (2D) plots

multi-layer 3D volumetric rendering

3D point clouds

external tool plug-ins (e.g. WorldWide Telescope)

spectrum (profile) display/analysis tools

2D plots from 3D data, in user-defined e.g. non-orthogonal coordinates

Open Session | Export Session | Import Data | Export Data/Subsets | Link Data | x² Arithmetic attributes | Active Subset: Perseus Only | Mode: | Terminal | Preferences

Data Collection

- Alves2020_Nat_678_237_MW3D[HDU1]
- Bialy2021_ApJL_919_L5_MW3D[HDU1]
- Pelgrims2020_A+A_636_A17_lmax10_MW...
- ONeill_Homology_Voids
- 2D_Bubbles_Multi_MW3D[HDU1]
- ONeill_LocalBubbleBfield
- Lallement2019_A+A_625_A135_solar_nei...
- ONeill_LocalBubble_lmax6
- Cahlon2023_Clouds[HDU1]

Subsets

- Perseus Only

Plot Layers

Alves-Bialy-Foley-O'Neill-Zucker | 3D views

3D view, note *selection*

2D (Sky) WWT

glue
multidimensional data exploration
glueviz.org

The 3D visualizations are just part of a much richer environment…

Linkable Interactive Visualization Exploration (LIVE) Environments

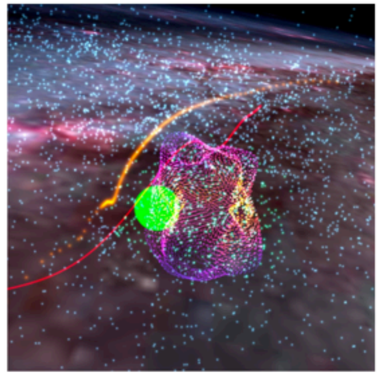
LIVE

What is LIVE?

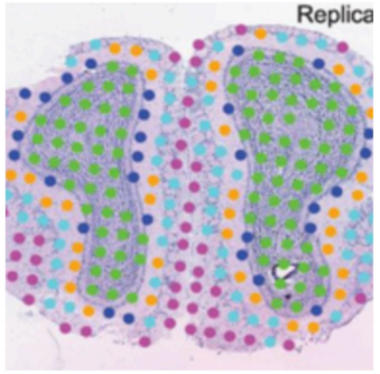
LIVE lets anyone build
"Linkable Interactive Visualization and Exploration" Environments.

LIVE is free, open-source, and helps with shared data and visualization challenges across astronomy (LIVE Astro), biology (LIVE Bio) and GIS (LIVE GIS).

As LIVE's infrastructure is being built, collaborators are ensuring its utility across Astronomy, Biology, and GIS by pursuing LIVE's science demonstration projects.



LIVE Astro



LIVE Bio

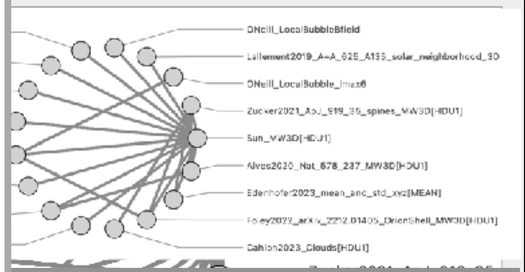
LIVE GIS

command-line terminal

ou can type any valid Python code here, and you
fined variables:

Link Editor

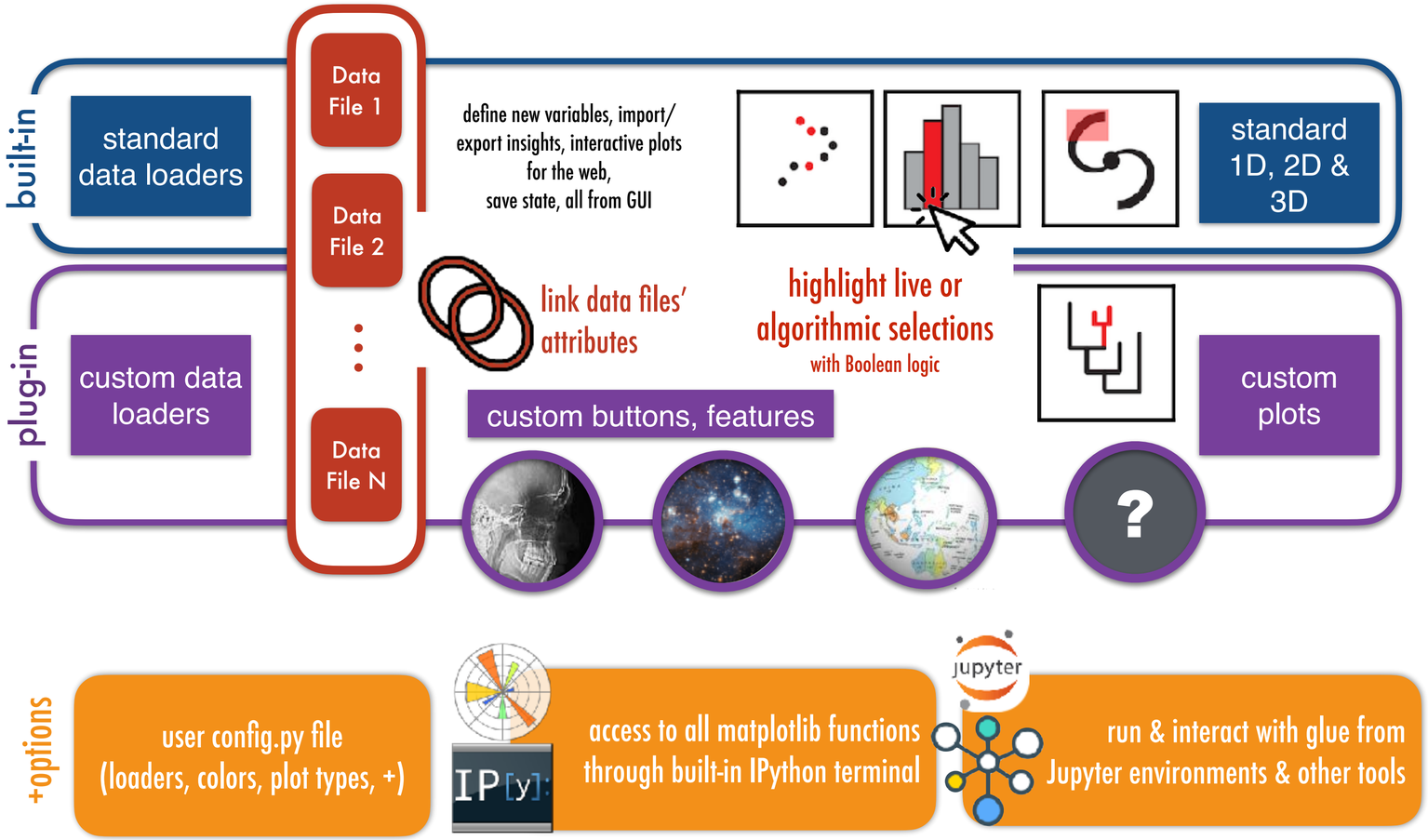
an existing connection to edit links. Selected datasets are shown
rs show directly and indirectly linked (blue) and inaccessible



GUI data set link editor

handout

glue
multidimensional data exploration
glueviz.org



learn more about glue

glueviz.org

Intro Video

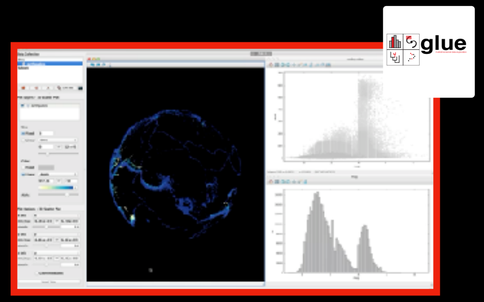
to combine glue with your favorite tools, in-browser, see LIVE

LIVE Environments

Exploratory Data Analysis (EDA)

Past

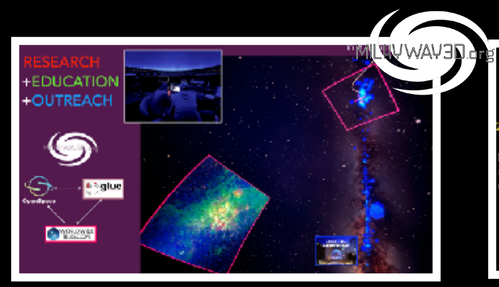
EDA in glue

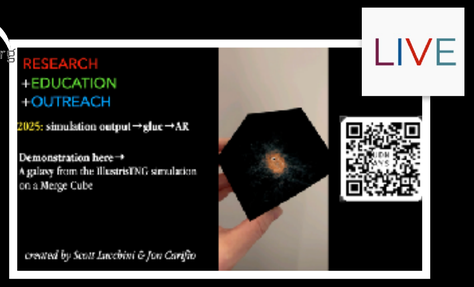


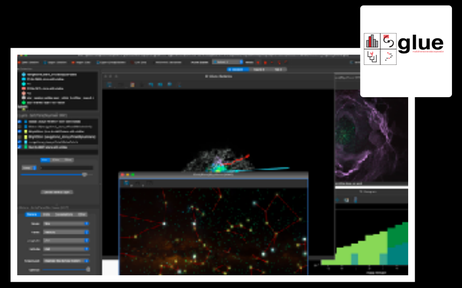
Present

++tools ++modalities

(phones, AR, planetariums, etc.)

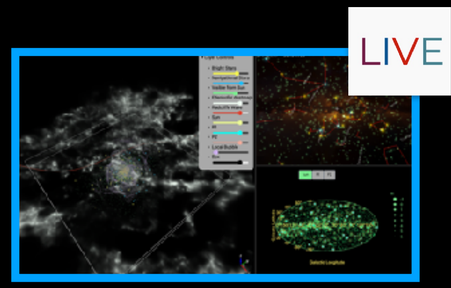


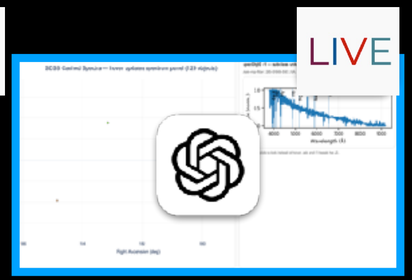


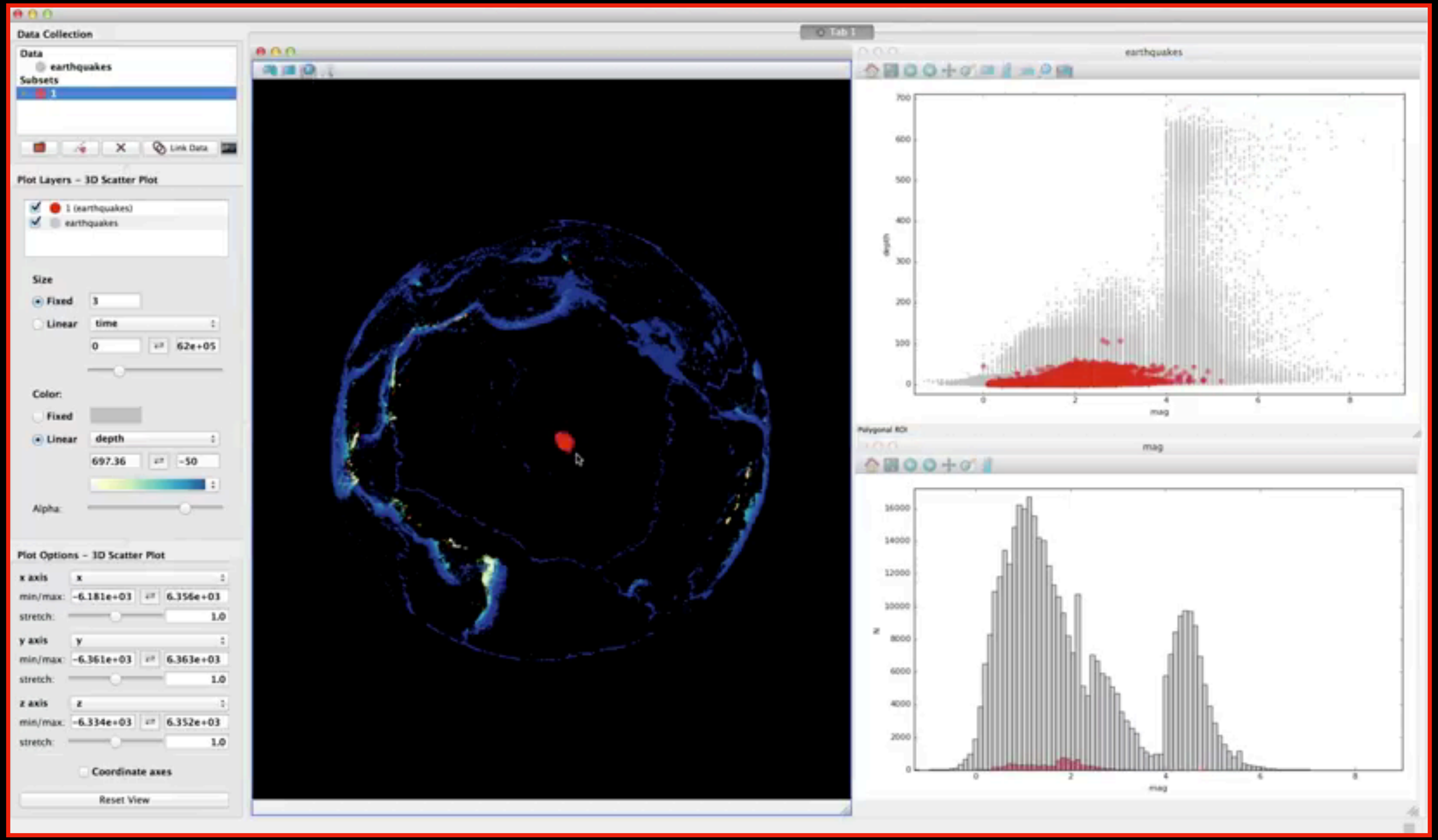


Future

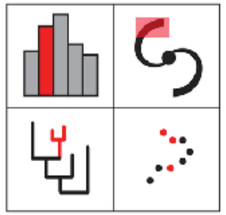
in browser, collaborative, AI assist







2015 video by Tom Robitaille, partner of Chris Beaumont in creating “glue 1.0”

What is  glue ?

multidimensional data exploration

It’s not an acronym.

It is open-source software that
glues data,
glues graphs &
glues tools.

data

numbers (tables, arrays, spreadsheets)

images & maps (FITS, JPEG, GIS and more)

data cubes (3D, 4D, and more)

data files' common attributes are glued

avoiding the need to merge data files

“graphs”

common statistical graphics

(scatterplots, histograms, tables, curves, overlays)

maps & images

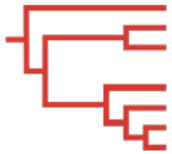
(greyscale, color, contours, layer control…)

3D displays

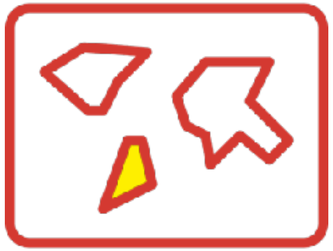
(scatter plots, volumetric rendering, sliders…)

specialized & custom charts

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



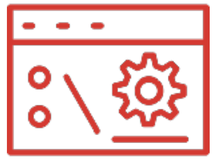
selections propagate across all graphs



for real-time data exploration & insight

tools

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



web services (across domains)

command-line (built-in terminal, scriptable)

for easy customization

glues data,
glues graphs &
glues tools.

glueviz.org

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.

glueviz.org

supported by

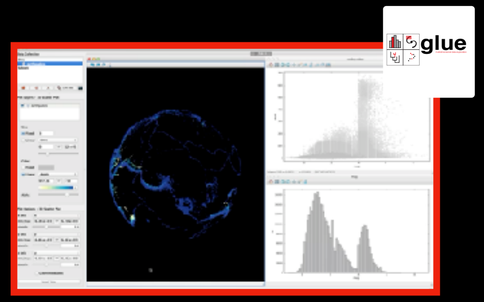
glue solutions inc.

GORDON AND BETTY MOORE FOUNDATION

Exploratory Data Analysis (EDA)

Past

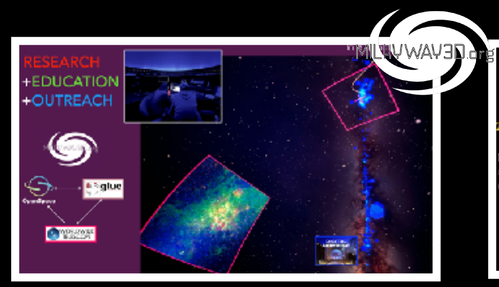
EDA in glue

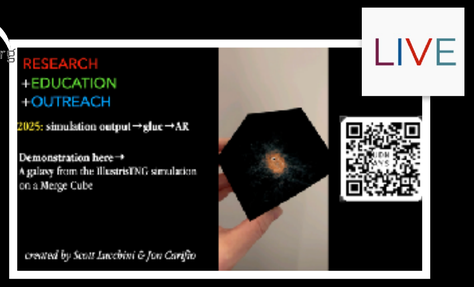


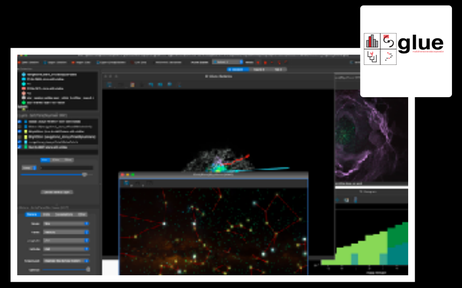
Present

++tools ++modalities

(phones, AR, planetariums, etc.)

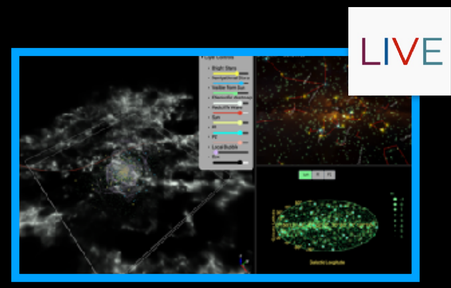


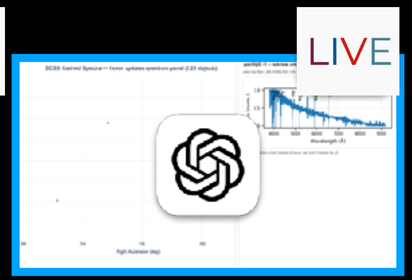




Future

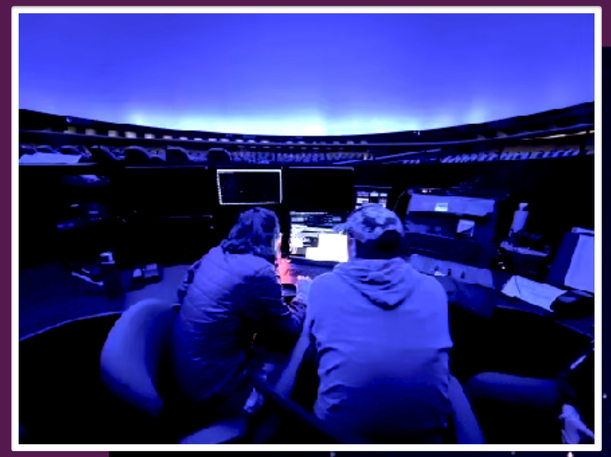
in browser, collaborative, AI assist





in a planetarium

RESEARCH
+EDUCATION
+OUTREACH

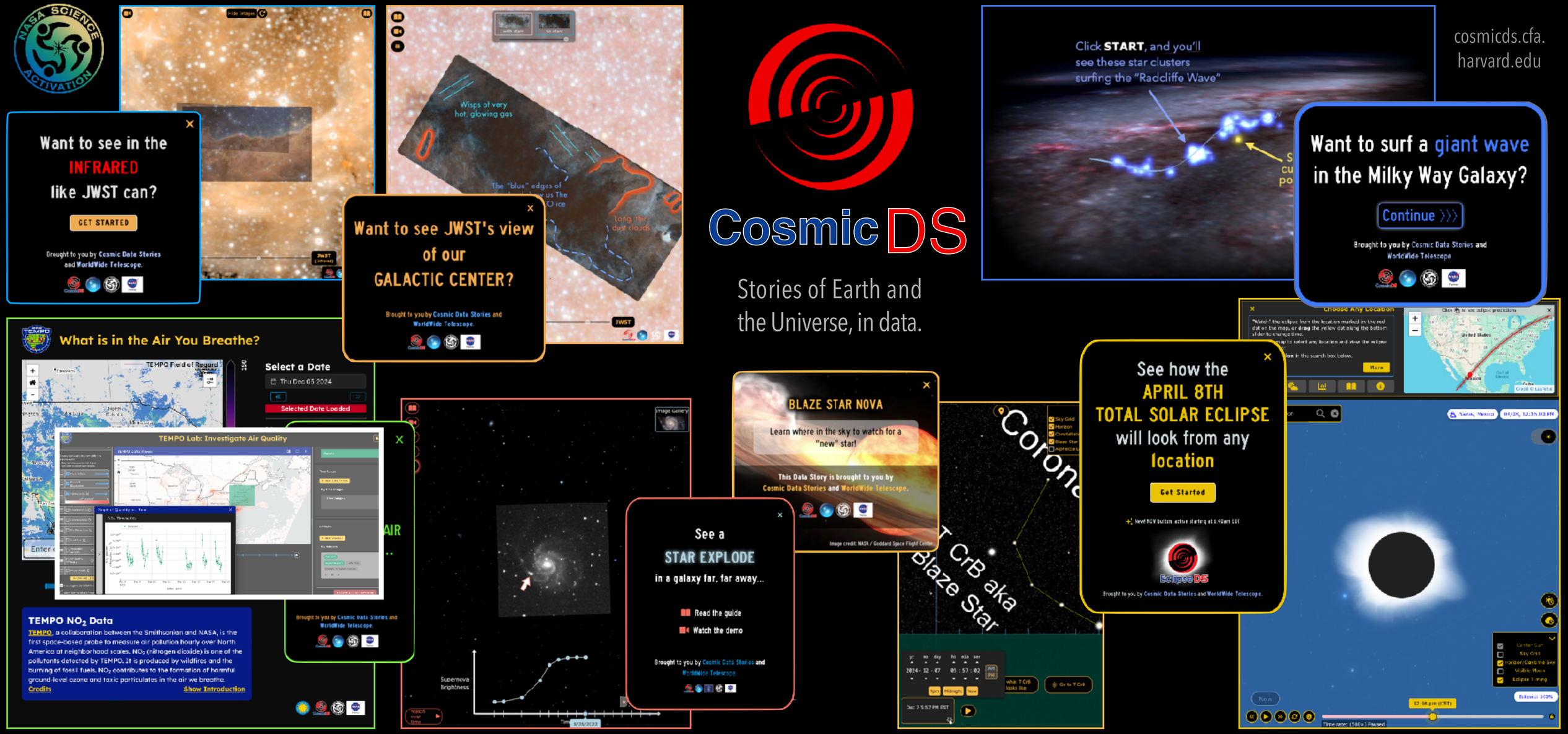


MILKYWAY3D.org

OpenSpace

glue

WORLDWIDE TELESCOPE

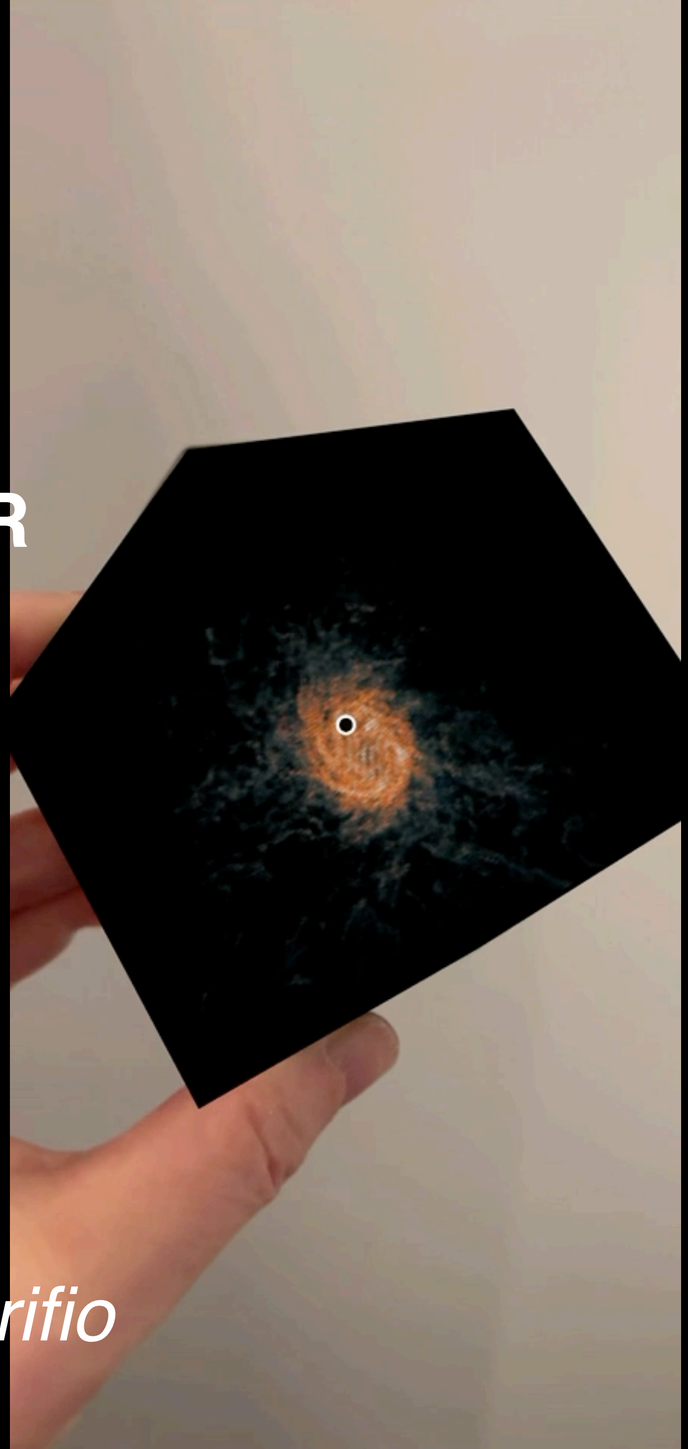


Ask CosmicDS team (Pat Udomprasert, John Lewis, Jon Carifo et al.) for more information

RESEARCH
+EDUCATION
+OUTREACH

2025: simulation output→glue→AR

Demonstration here→
A galaxy from the
IllustrisTNG simulation on a
Merge Cube



created by Scott Lucchini & Jon Carifio

Exploratory Data Analysis (EDA)

Past

EDA in glue

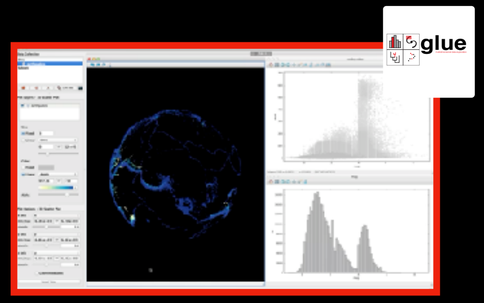
Present

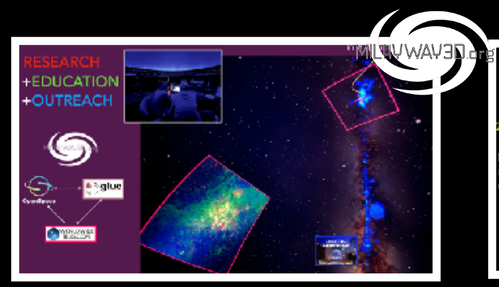
++tools ++modalities

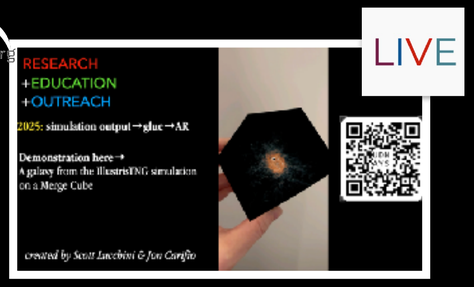
(phones, AR, planetariums, etc.)

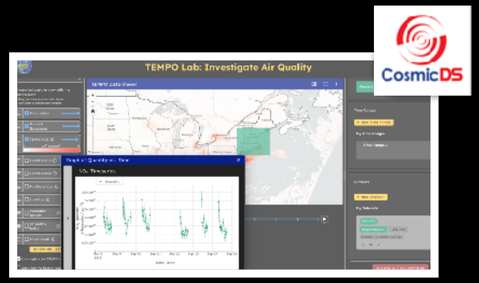
Future

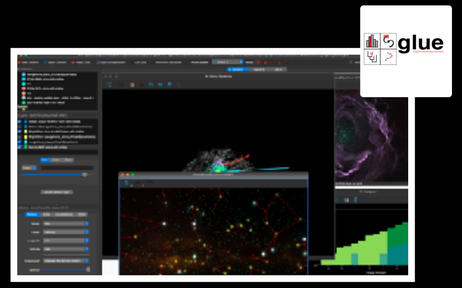
in browser, collaborative, AI assist

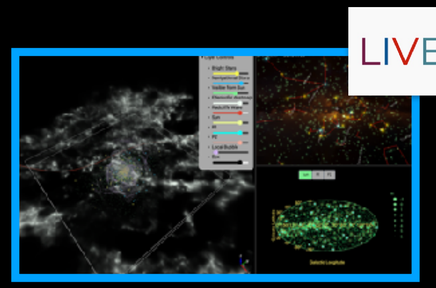


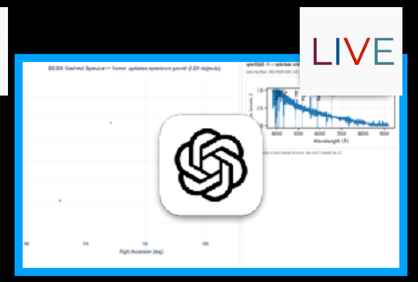




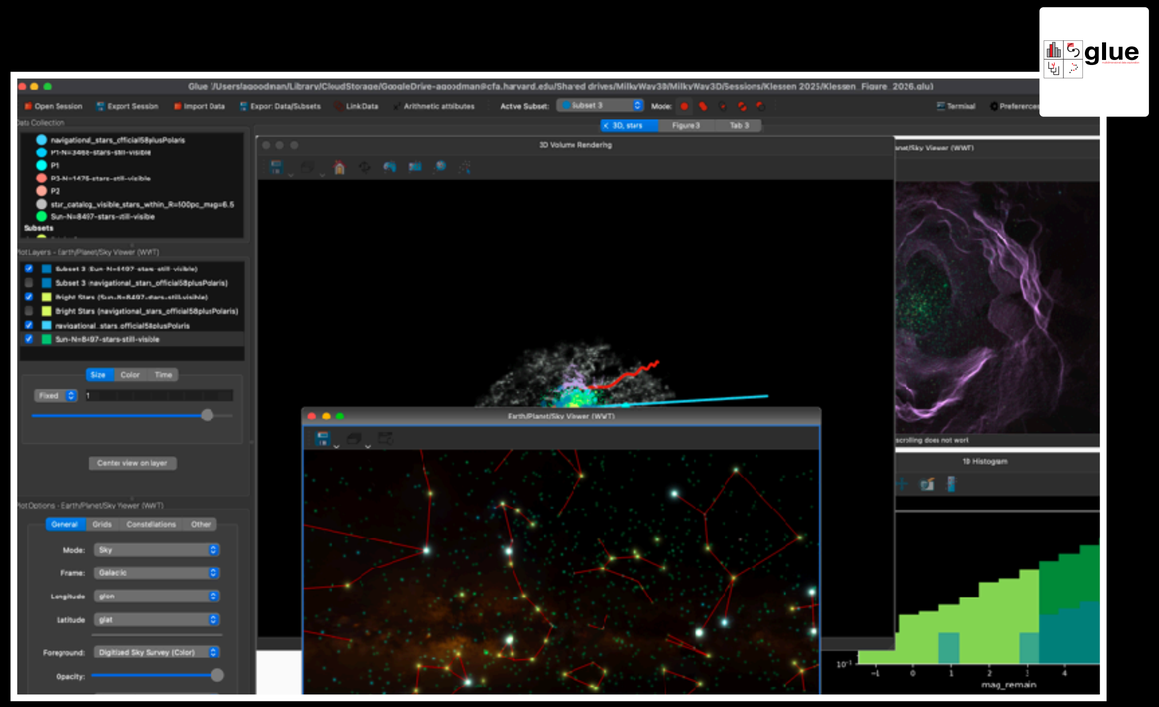


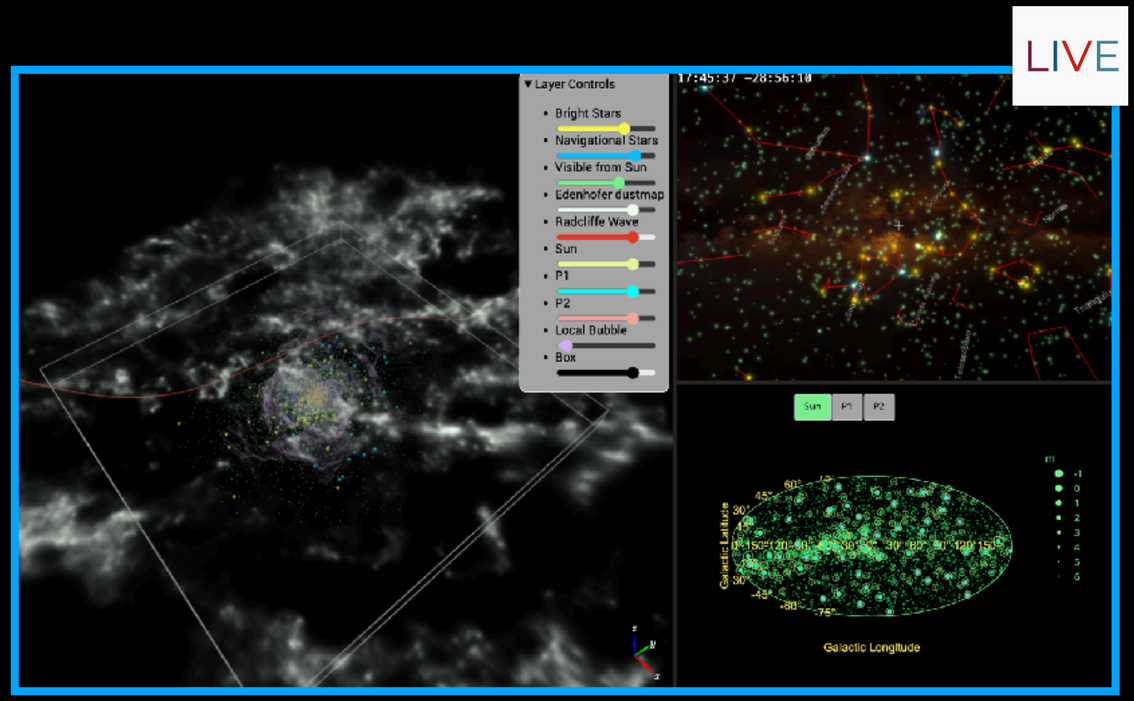






“From Exploration to Explanation, and back”—Interactive Figur





for Klessen et al. 2026 (forthcoming)

Exploratory Data Analysis (EDA)

Past

EDA in glue

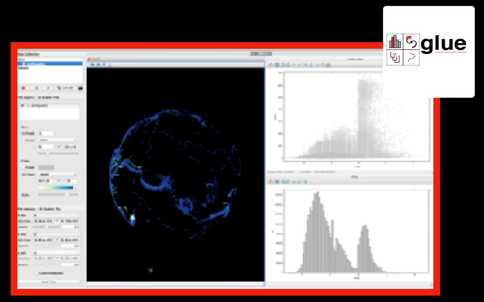
Present

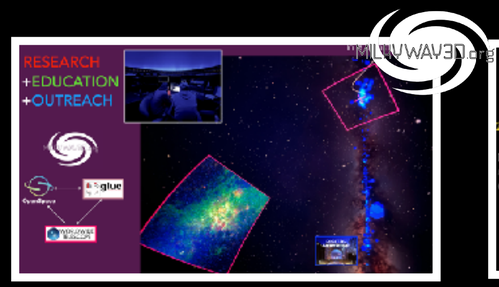
++tools ++modalities

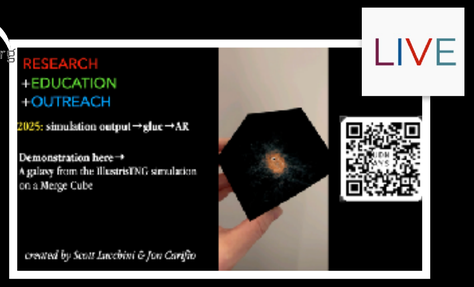
(phones, AR, planetariums, etc.)

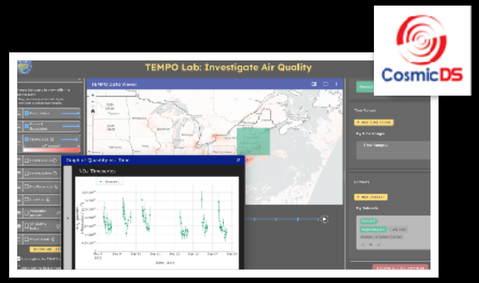
Future

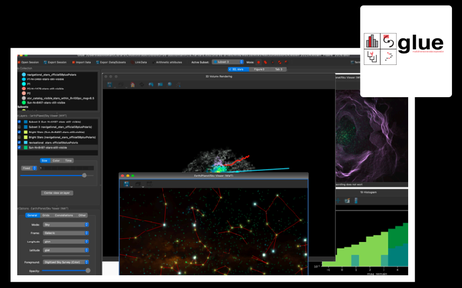
in browser, collaborative, AI assist

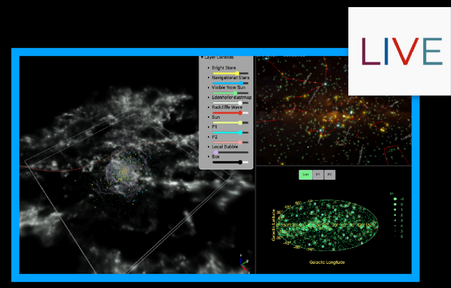


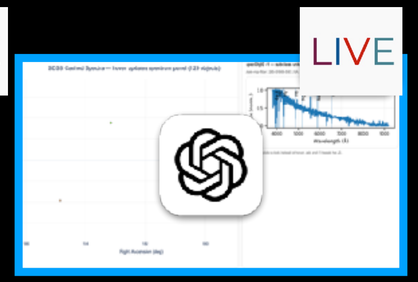












[live demo]

Exploratory Data Analysis (EDA)

What is it, and why does it matter in the age of big surveys and AI?

In our future scientific investigations, computers will be able to do *almost* anything we ask.

The trick will be *knowing what to ask.*

Human + AI “partnerships” via Exploratory Data Analysis (EDA)

EDA is a conversation with data — AI will join in.

INFORMATION

Data

Simulation

Theory

Exploratory Data Analysis

IDEAS

Hmm… that’s weird?!

New hypothesis

Present
I can help with: data retrieval, formatting, coding, optimization, making specific visualizations…

Future
Can I join you up there?
I think I see a pattern. Want to test if it’s real or an artifact?

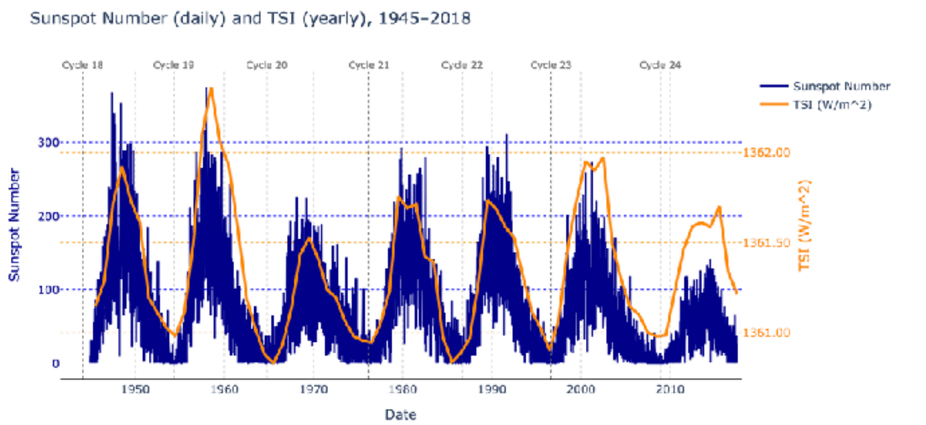
For discussion…

Jdaviz

CosmicDS

Home About Data Stories Resources

And, finally, here's a Sunspot Cycle Data Exploration, Interactive Graphic, created by visualization expert using only ChatGPT, as a demonstration.



This interactive plot shows the daily Sunspot number (in blue) and Total Solar Irradiance (in orange) for 1945-2018, with lines indicating the beginning of each solar cycle.

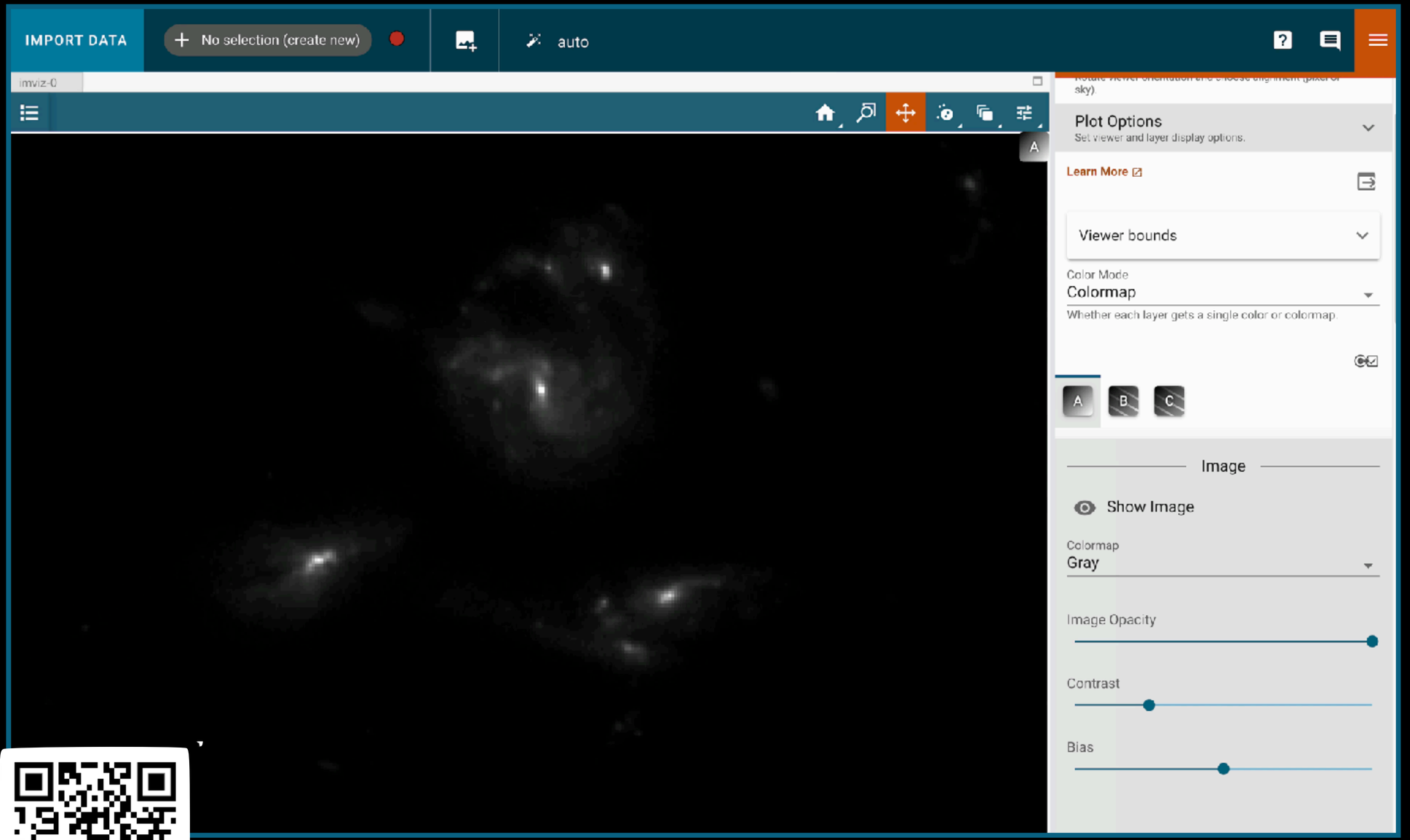
ask
about your
favorite tool

Table A									
Software	**Function**	**Browser (B) Desktop (D) Python anywhere (Py)**	**(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**	**Coordinating Contacts**	**License Type**
dask	Data Analysis	D, B		1	scales Python data analysis with parallel, out-of-core computing	1	github.com/dask/	Dask community	BSD 3
Astropy	Data Analysis	Py		1	core python library used for astronomy functions across tools	1	github.com/astropy	Astropy community	BSD 3
Jdaviz	Domain tool	B		1	visualization package built on glupyter for JWST; upcoming version will run on Solara	1	github.com/spacetelescope/jdaviz	NASA/STScI	BSD 3
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	MIT
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/WorldWideTelescope/	SAO/Harvard	MIT
glue	EDA	D		1	linked-view data visualization, also connects software packages via plug-ins (precursor to LIVE)	1	github.com/glue-viz	glue team at SAO/Harvard	BSD 3
glupyter	EDA	B		1	Jupyter-based version of glue (see glue)	1	github.com/glue-viz/glue-jupyter	glue team at SAO/Harvard	BSD 3
glue-JupyterLab	EDA	B		1	glue extension for JupyterLab	1	github.com/QuantStack/glue-jupyterlab	Quantstack	BSD 3
JupyterLab/ Notebook	Framework	B		1	frameworks for running Julia, python, and R in the browser	1	github.com/jupyterlab/jupyterlab	Project Jupyter Community/ Quantstack team	BSD 3
Solara	Framework	B		1	Python framework for reactive apps, used in CosmicDS and new version of Jdaviz for Roman Space Telescope	1	github.com/widgetti/solara	Widgetti community	MIT
ipyAladin	Sky Viewer	Py		1	bridge between Jupyter and Aladin Lite	1	github.com/cds-astro/ipyaladin	Strasbourg Astronomical Observatory	BSD 3
bqplot	Visualization	B		1	2-D visualization system for Jupyter, based on the constructs of the Grammar of Graphics	1	github.com/bqplot/	Projec Jupyter community /QuantStack team	Apache 2.0
matplotlib	Visualization	Py		1	primary plotting library used in glue and Jupyter	1	github.com/matplotlib/	(fiscally NumFOCUS)	PSF (BSD compatible)
Plotly	Visualization	Py		1	general-purpose highly-customizable visualization tool with excellent support for interactivity	1	github.com/plotly/plotly.py	Plotly community	MIT
JupyterAI	AI enhancement	Py		1	Generative AI tools seamlessly integrated into Jupyter	2	https://github.com/jupyterlab/jupyter-ai	Project Jupyter community	BSD 3
Dataverse	Data archive	B		1	field-agnostic data-hosting platform, assigns	2	github.com/IQSS/dataverse	IQSS/Harvard	Apache 2.0

Jdaviz

Data analysis and visualization of astronomical images and spectra in Python.

Works in Jupyter Notebooks, as a standalone application, and embedded in webpages.

Links data & views using glue in the browser "glupyter."

video, created by C. Pacifici, shows JADES survey (NIRSpec GTO) in Jdaviz