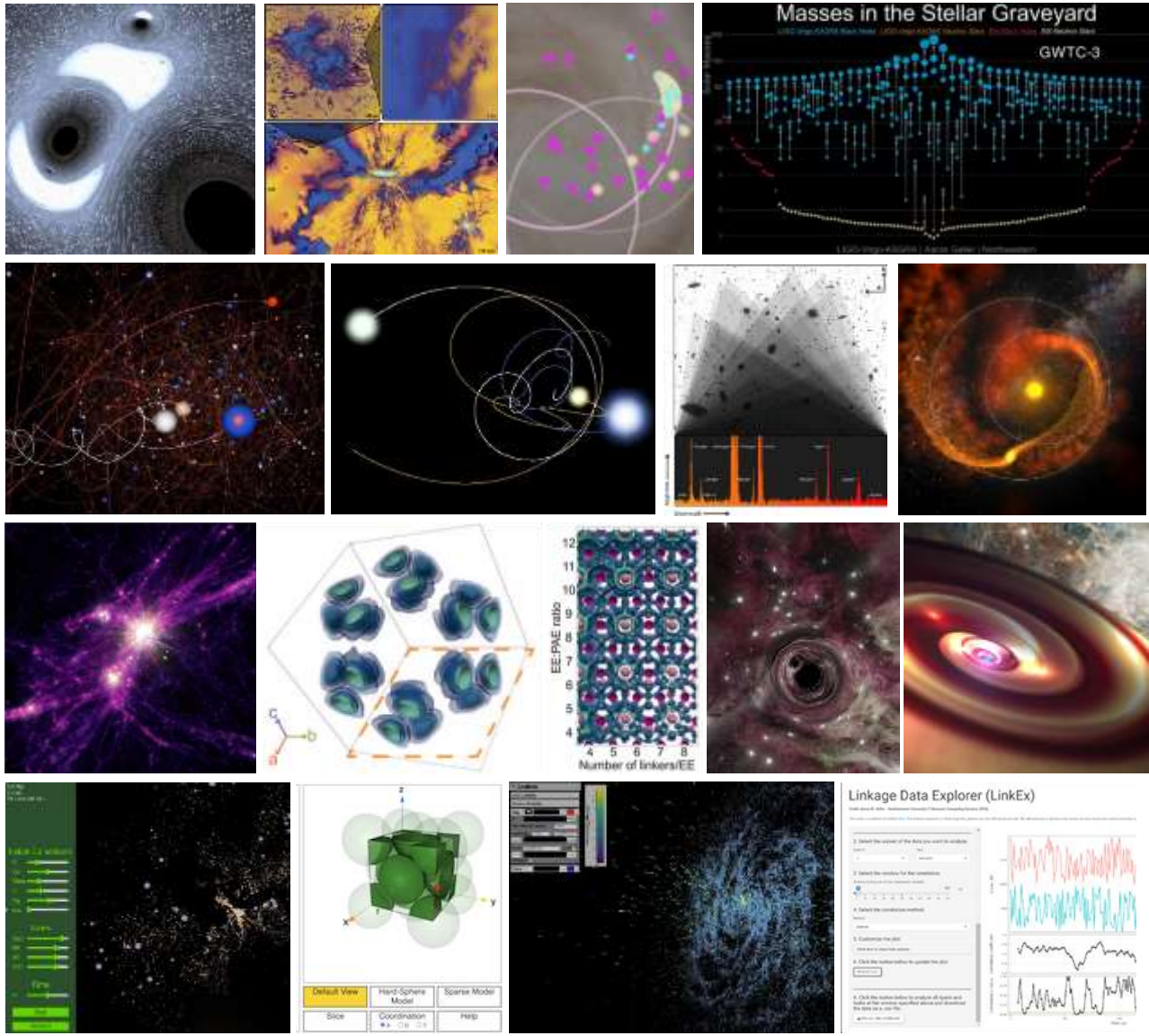
Why/How/When to Visualize Your Research Results

Aaron M. Geller

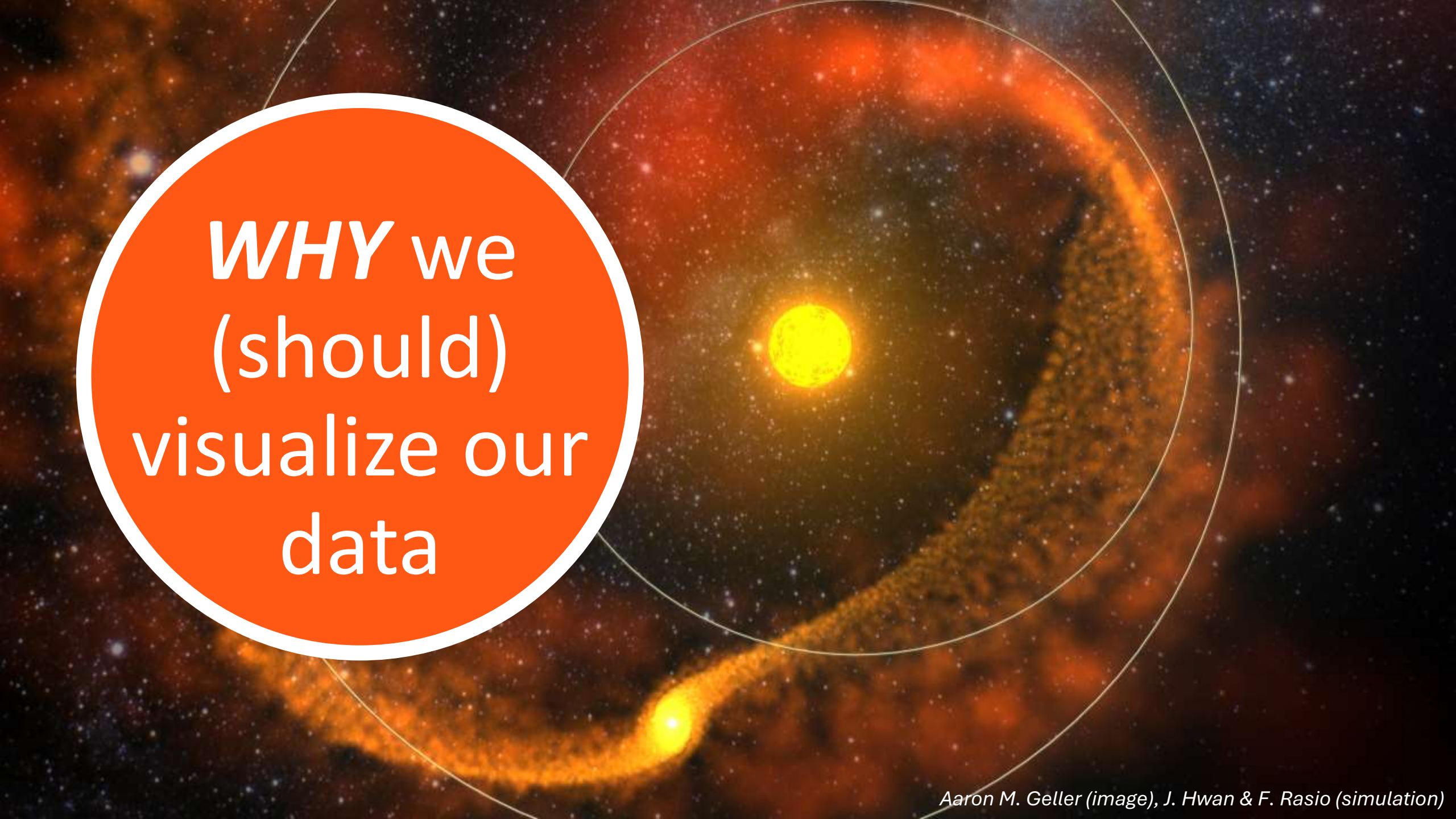
Senior Data Visualization Specialist
Research Computing and Data Services

Research Associate Professor
Center for Interdisciplinary Exploration and Research in Astrophysics (CIERA) and Department of Physics & Astronomy



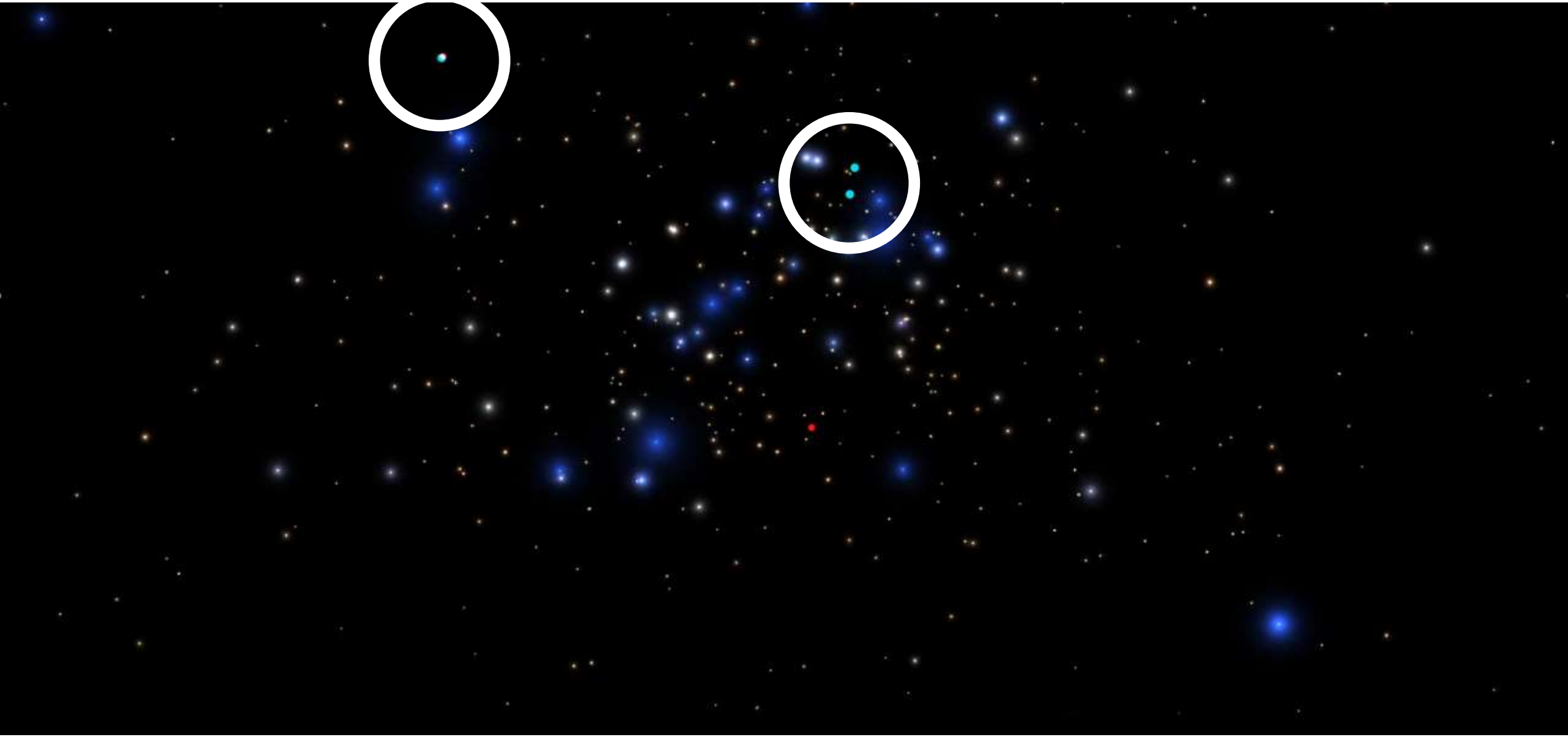
CIERA
CELEBRATING
15 YEARS

Northwestern
INFORMATION TECHNOLOGY
RESEARCH COMPUTING AND DATA SERVICES



Aaron M. Geller (image), J. Hwan & F. Rasio (simulation)

Why : Scientific discovery!



Why : Bug checking

Why: Data Expl-oration / -anation

	Exploration *Help YOU learn about your data*	Explanation *Help OTHERS learn about your data*
Audience	You and your collaborators	Pick one
Number of Visualizations	Many	Probably 1 per dataset
Visualization Message	Unknown	Why include the visualization?
Formatting	Not important	Important, possibly restricted

OC Binary Explorer

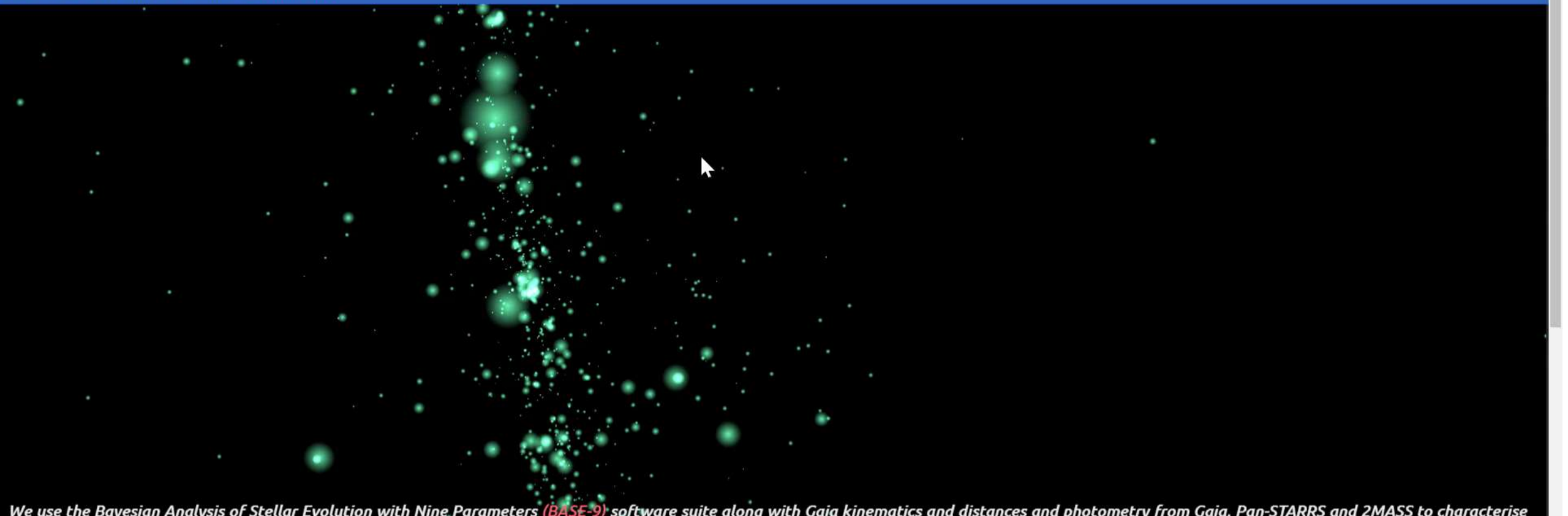
ocbinaryexplorer.ciera.northwestern.edu

Guest

Open Cluster Binary Explorer

Access photometric binary-star data in a collection of open clusters

An NSF-funded research project from PI Aaron M. Geller.



We use the Bayesian Analysis of Stellar Evolution with Nine Parameters (BASE-9) software suite along with Gaia kinematics and distances and photometry from Gaia, Pan-STARRS and 2MASS to characterise the binary-star populations in a collection of open clusters. For information about our methods and results, please see the Papers section at the bottom of this page. You can access these data by clicking on the button below.

The interactive visualization above shows the open cluster population as seem from Earth, with the clusters in this study highlighted in pink.

Click here to enter the *Interactive Data Explorer*

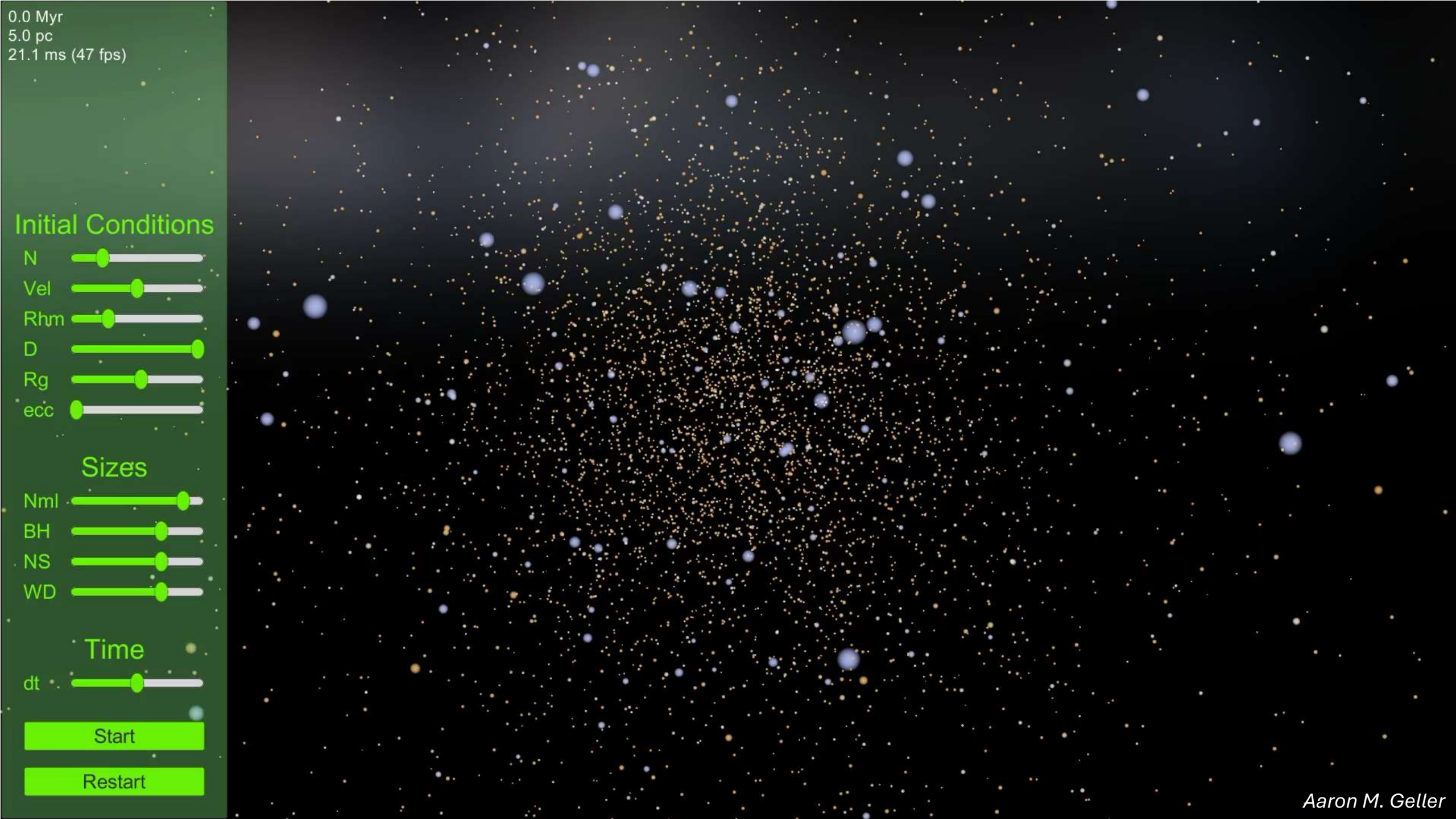
View, filter, sort, create, edit, and download data and plots

Unity

GPGPU Star Cluster Viz

Aaron Geller

https://github.com/ageller/UnityNbodyCluster


0.0 Myr
5.0 pc
21.1 ms (47 fps)
Initial Conditions
N
Vel
Rhm
D
Rg
ecc
Sizes
Nml
BH
NS
WD
Time
dt
Start
Restart
Aaron M. Geller

Firefly:

A WebGL tool to explore particle-based data

Aaron Geller / Alex Gurvich / Northwestern

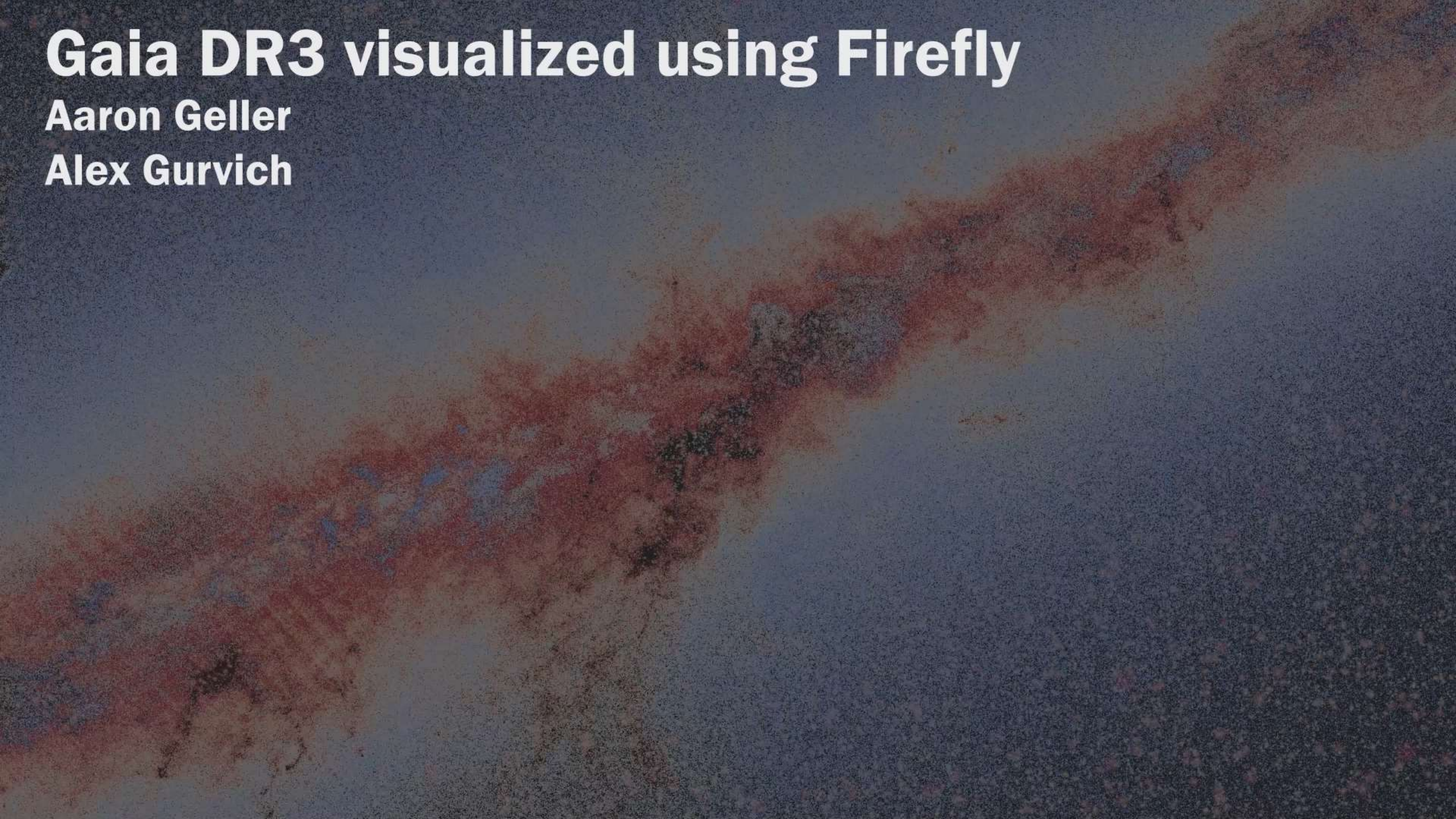
Instructions:

- Right-click and drag with the mouse to rotate your view.
- Use the mouse wheel to zoom in and out.
- Click the Controls bar on the top left to show/hide a user interface.
- Detailed instructions can be found on the Firefly GitHub page.
- **h**: toggles this help screen on and off.

Click to begin.

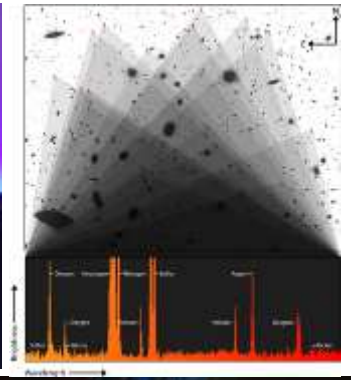
Gaia DR3 visualized using Firefly

Aaron Geller
Alex Gurvich

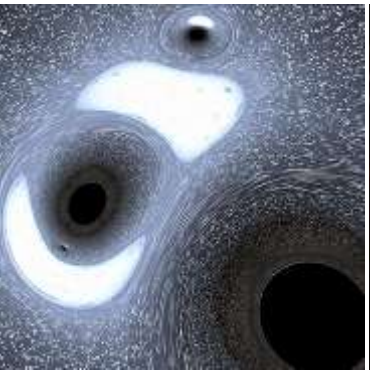
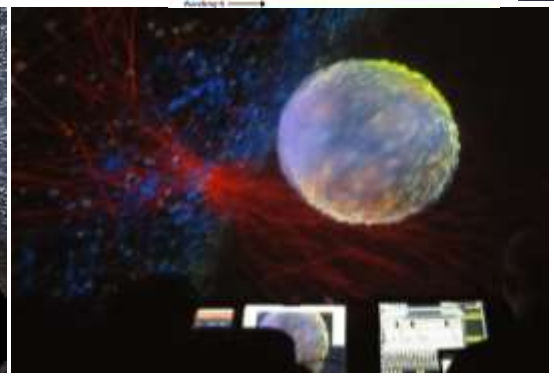
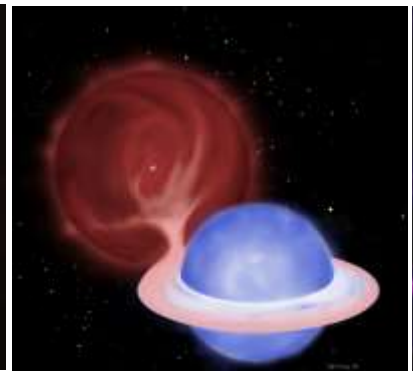
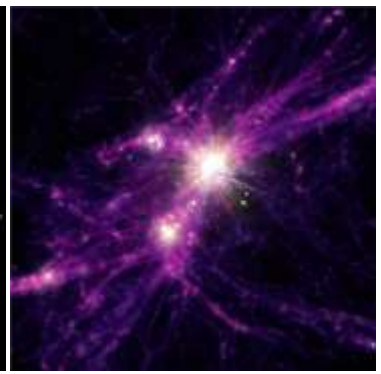


Why : Public Engagement

In person and online / press



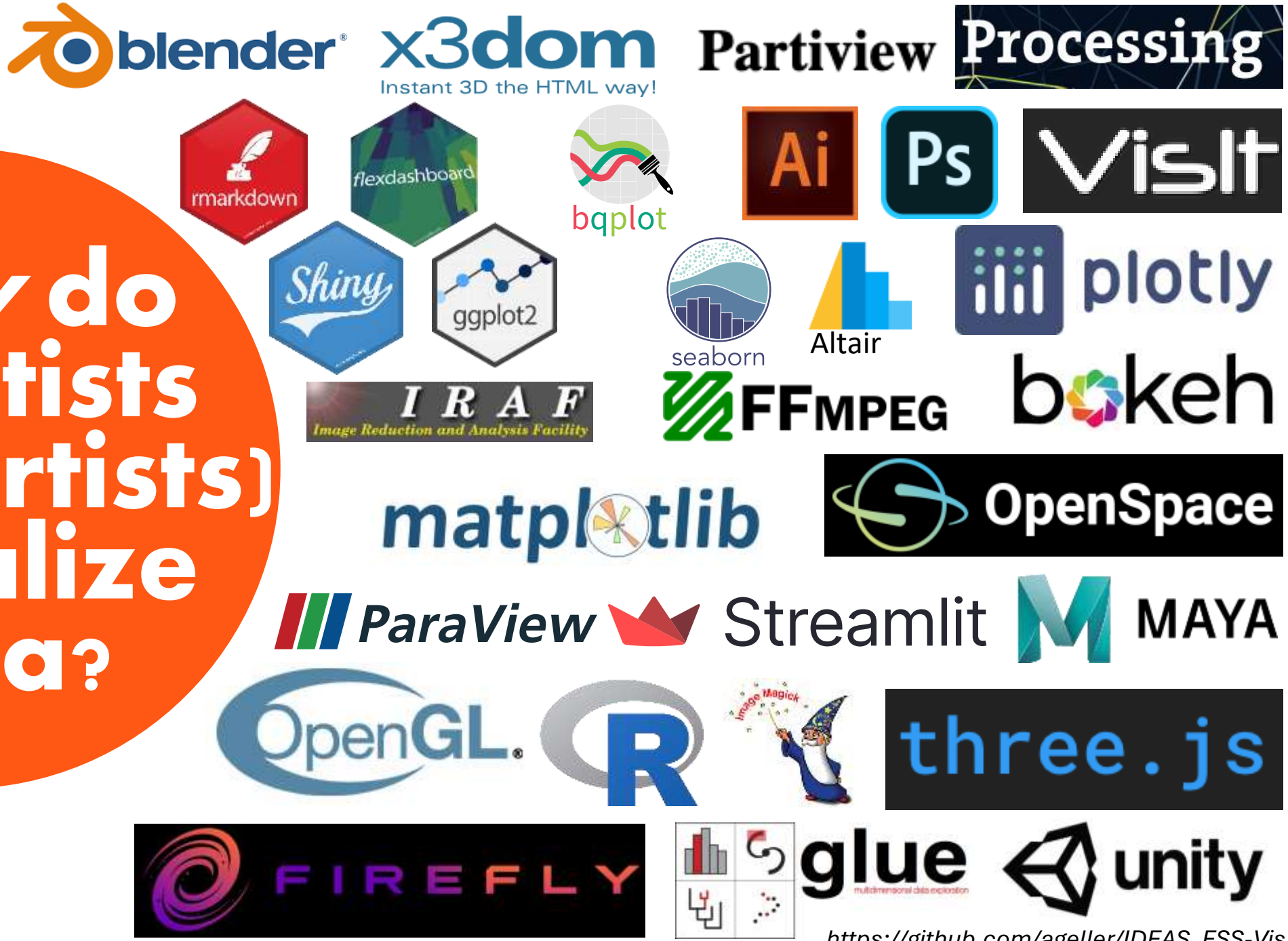









Masses in the Stellar Graveyard

How do scientists (and artists) visualize data?



https://github.com/ageller/IDEAS_FSS-Vis

How: Some General Suggestions

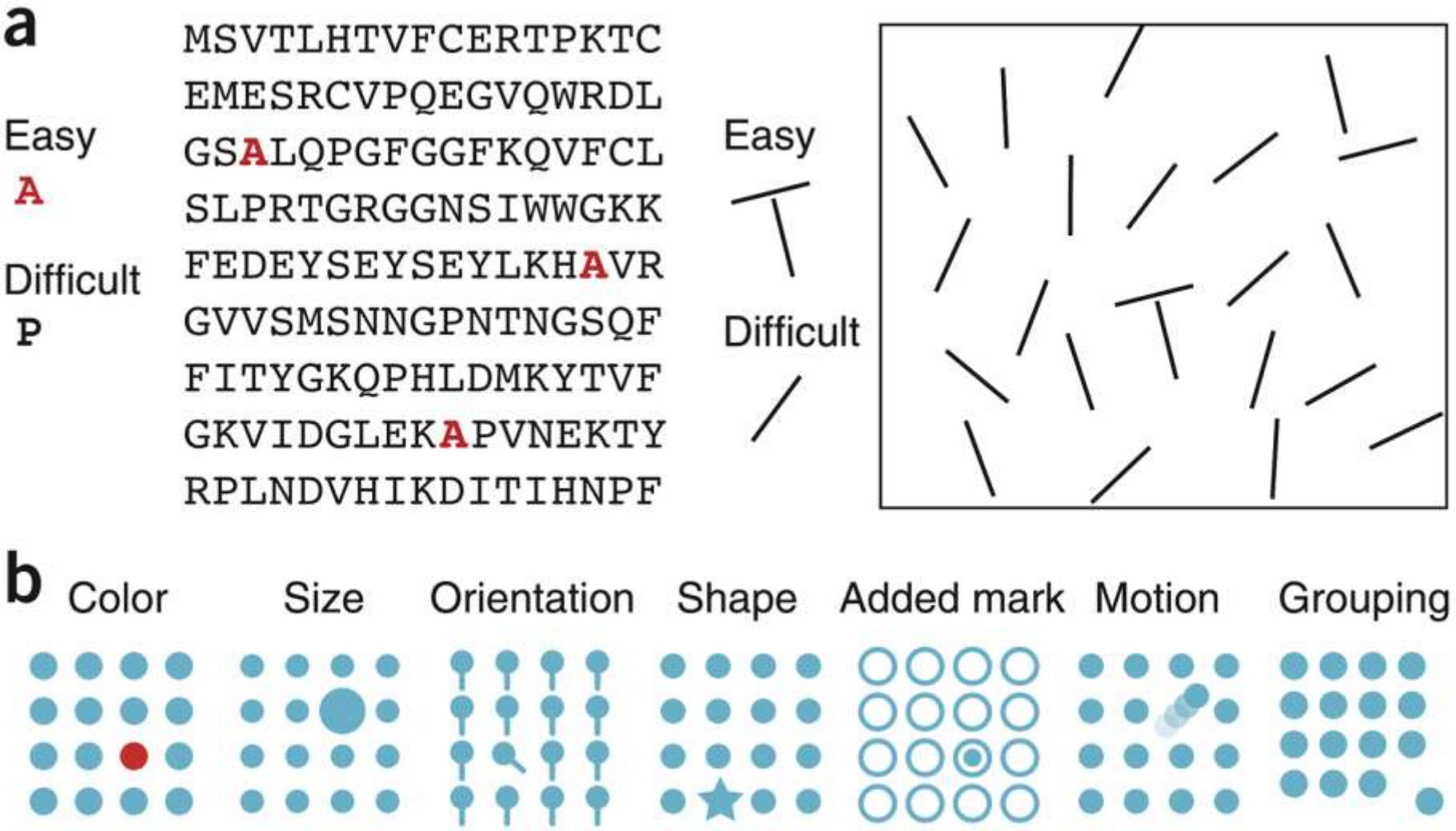
- **Focus on storytelling.**
- **Improve Data-to-Ink Ratio:** *"Is this [shape] necessary?"*
- **Ensure Visual Quality:** *"Is this geometry telling the truth?"*
- **Reduce Clutter:** *"Is this color choice or layout necessary?"*
- **Increase Efficiency:** *"Is it too hard or time consuming to read?"*
- **Consider Accessibility:** *"Is this colorblind safe? Is the font size large enough?"*
- **Organize and Guide:** *"Should I regroup my data? Can I add helpful text?"*

Edward Tufte's "Data to Ink Ratio"

https://www.darkhorseanalytics.com/blog/data-looks-better-naked

How : Salience

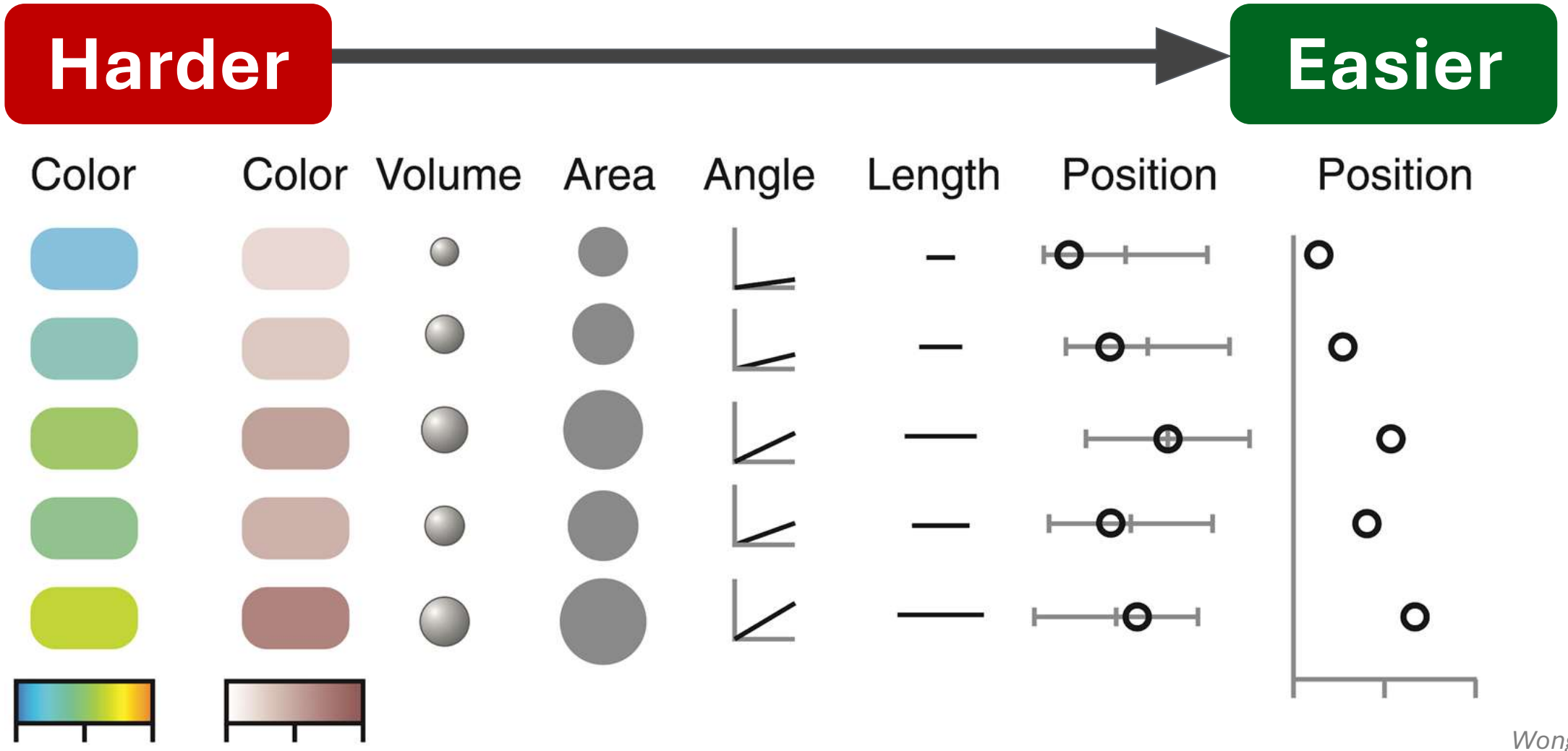
Guide your viewer to your result



Wong (2010)

How : Salience

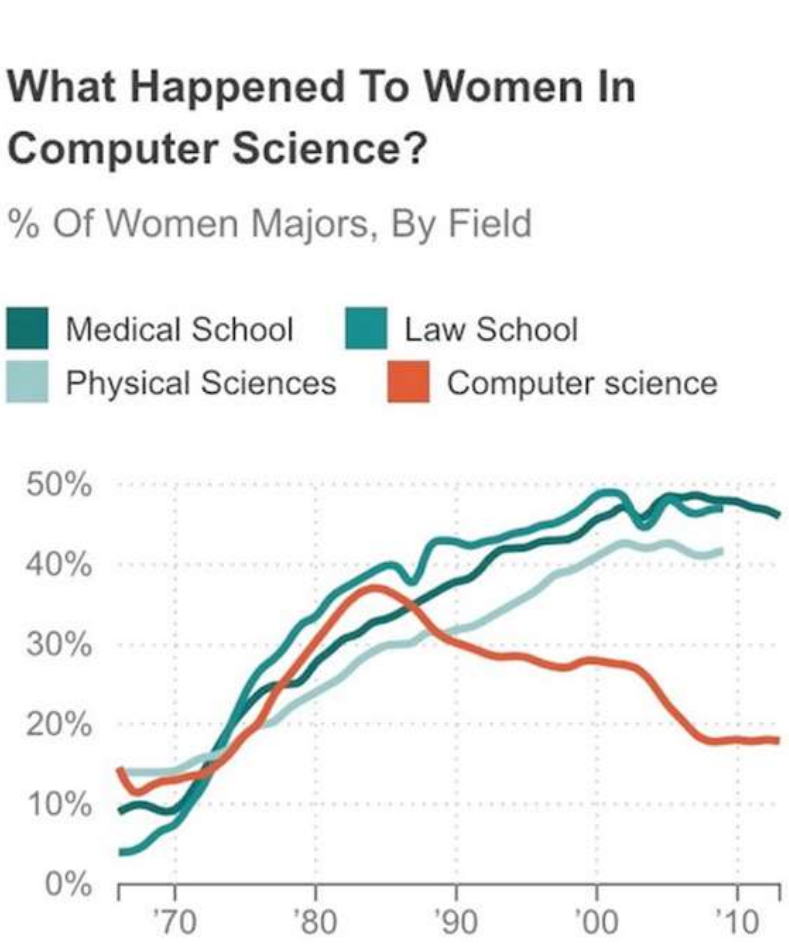
Use easy-to-estimate visual representations

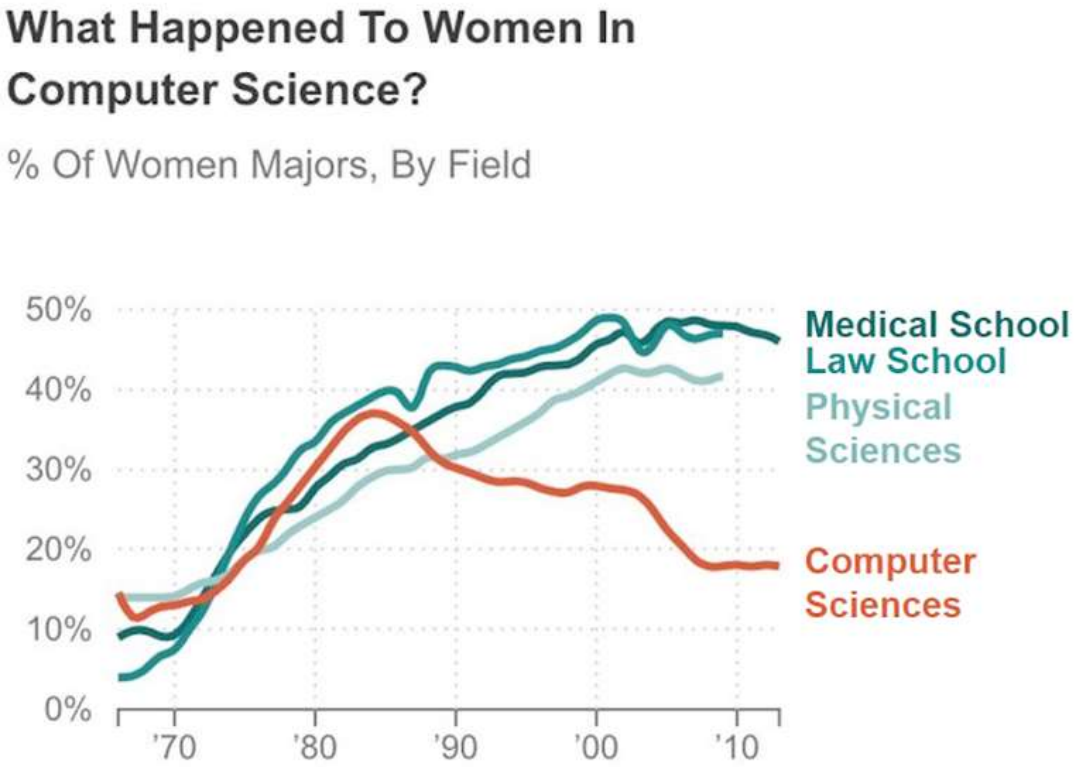


Wong (2010)

How : Salience

Try direct labels instead of legends; highlight the most important result

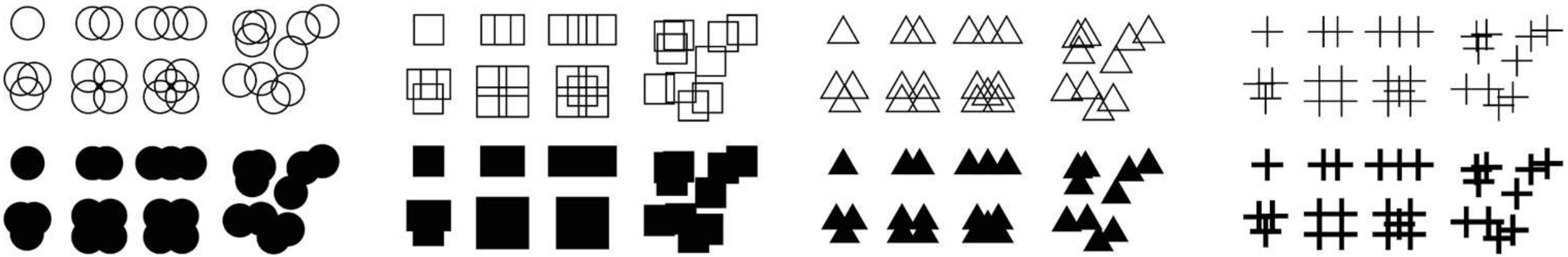




https://depictdatastudio.com/directly-labeling-line-graphs/

How : Symbols

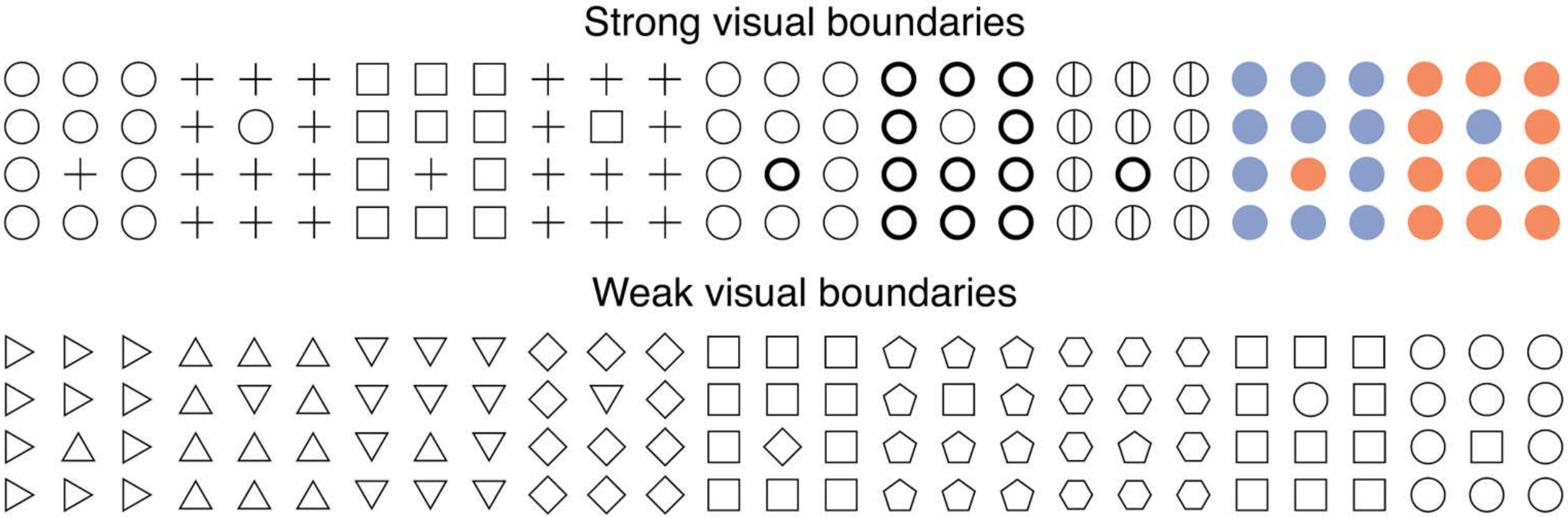
Open circles are the most flexible.



Krzywinski & Wong (2013)

How : Symbols

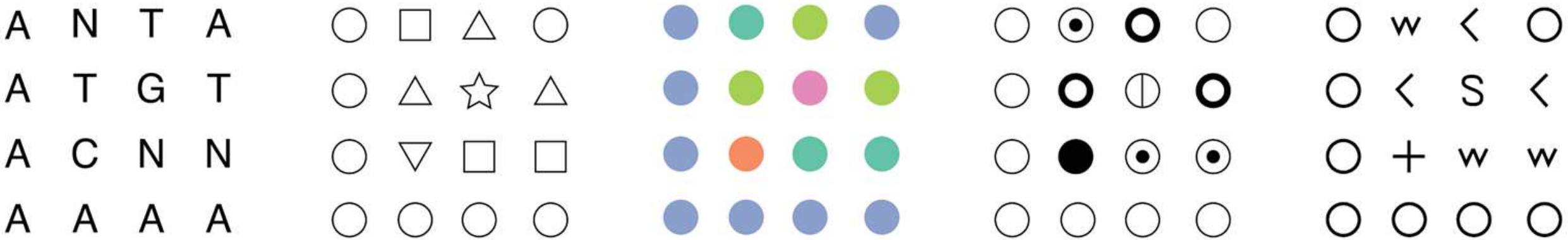
Form strong visual boundaries



Krzywinski & Wong (2013)

How : Symbols

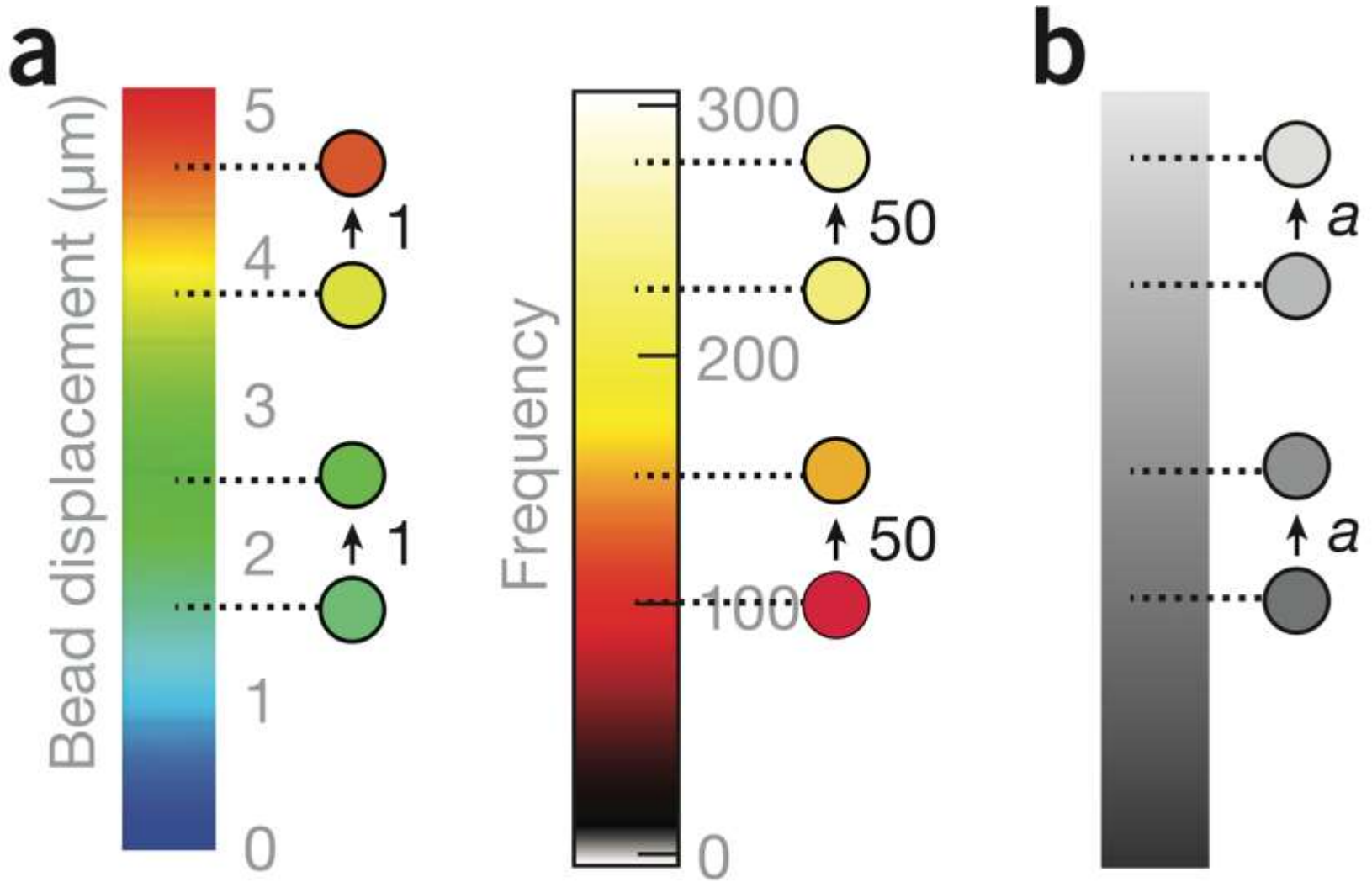
Form strong visual boundaries



Krzywinski & Wong (2013)

How : Colors

Choose colormaps wisely but note that color is not ideal for representing quantitative data.

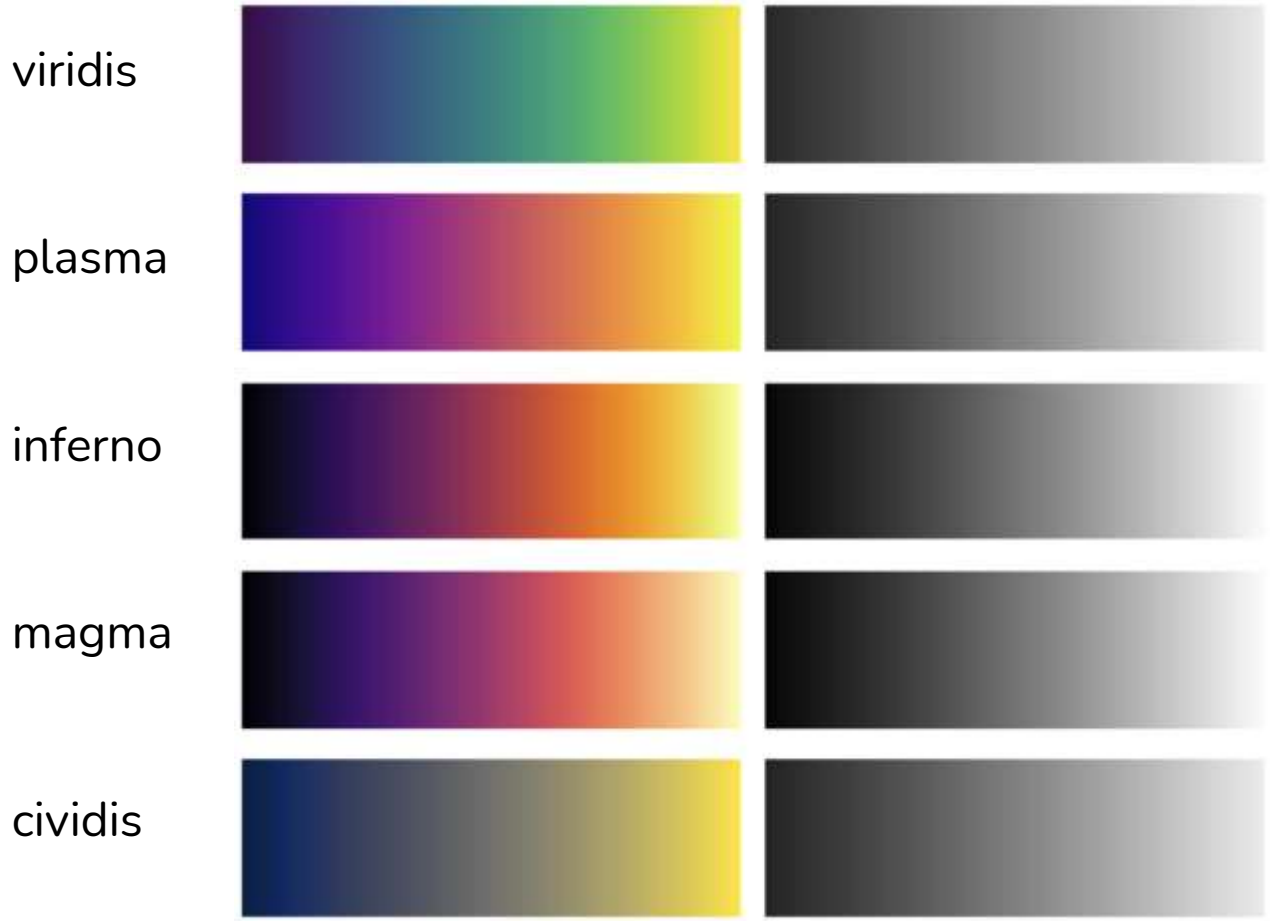


Wong (2011)

How : Colors

What does it look like in greyscale? Is it colorblind safe?

Perceptually uniform sequential colormaps

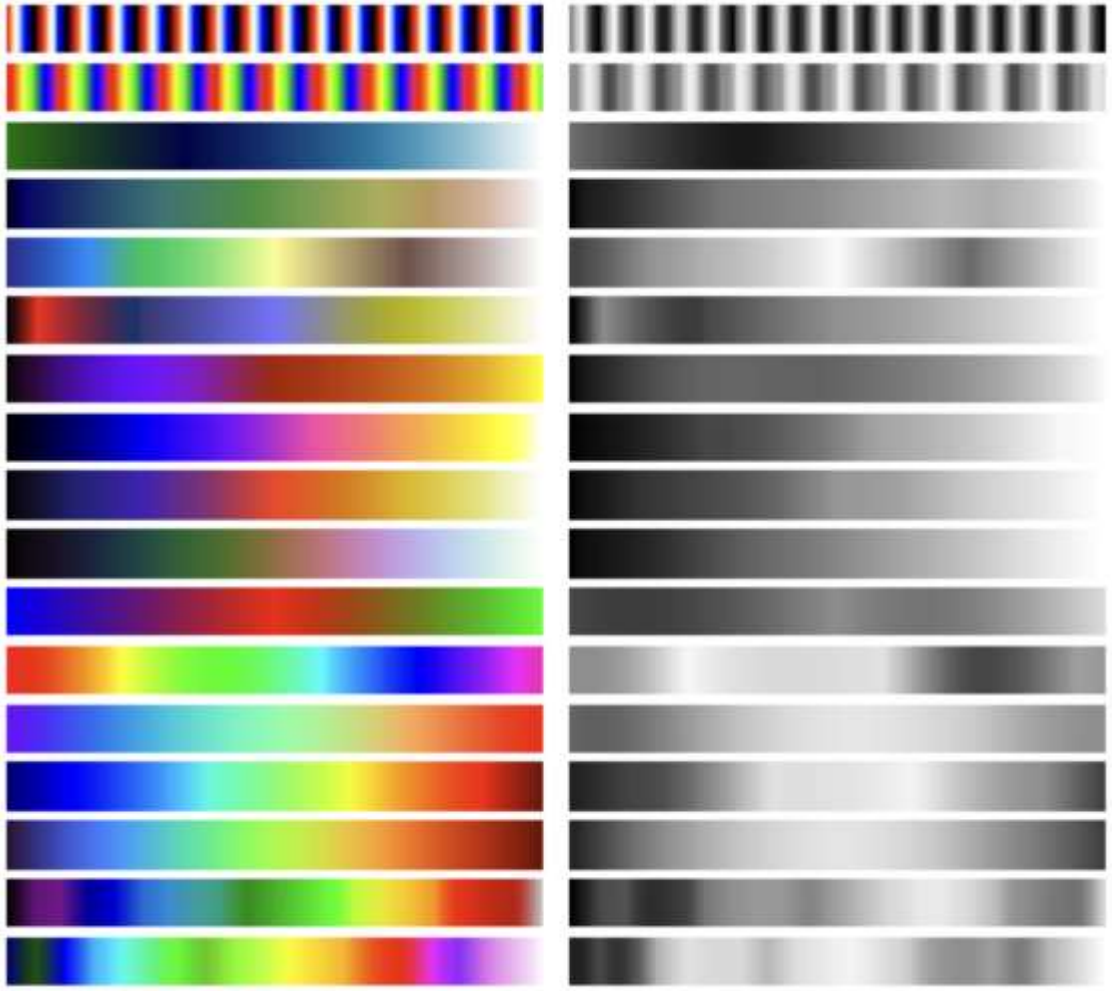


https://colorbrewer2.org/
https://coolors.co/
matplotlib
https://www.color-blindness.com/coblis-color-blindness-simulator/

How : Colors

What does it look like in greyscale? Is it colorblind safe?

There are LOTS of (often very bad) colormaps out there. Be careful!



matplotlib

How : Colors

Colors can have meaning (that differs between cultures).

Good

Bad

Banana

Sky

How : Colors

Colors can have meaning (that differs between cultures).

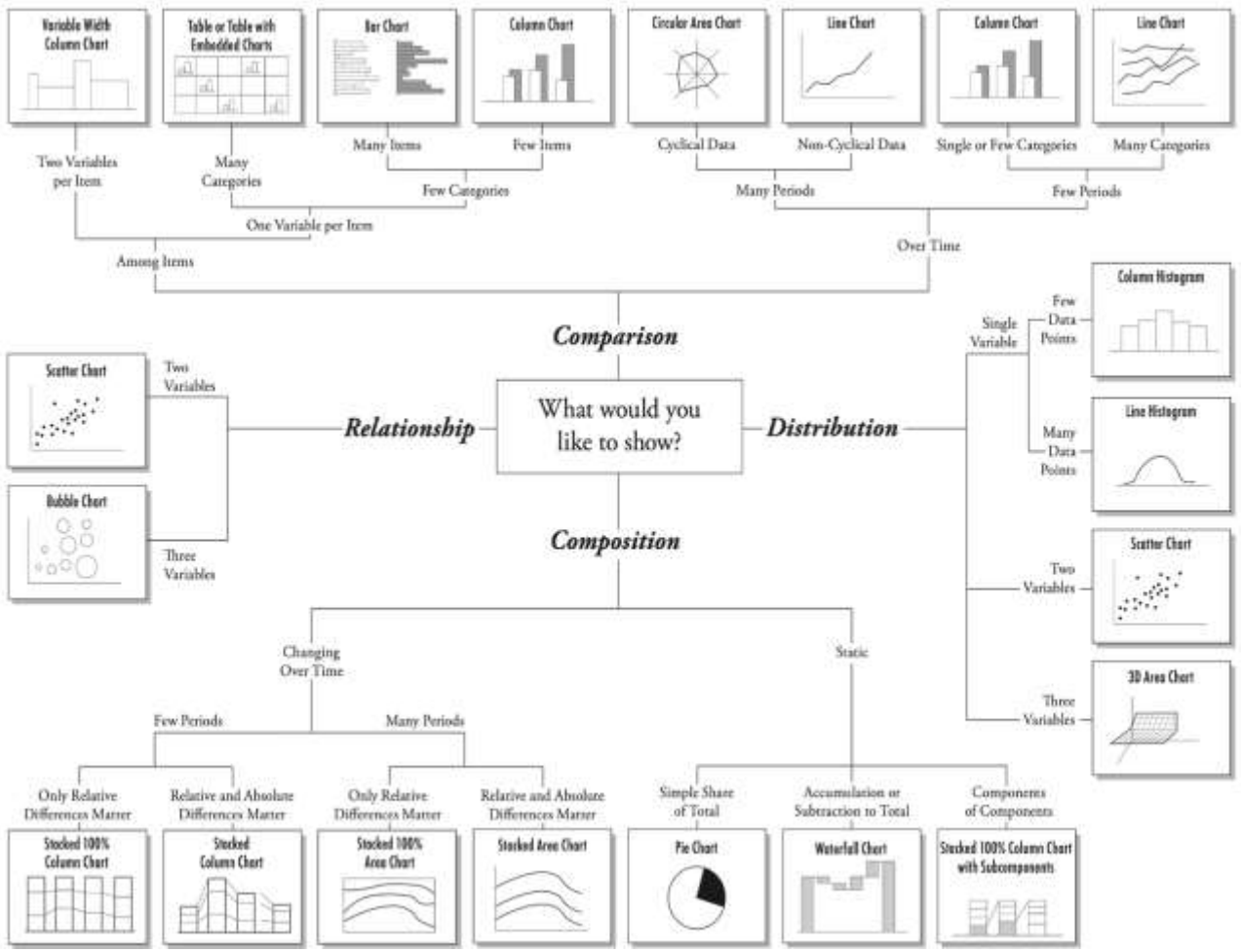
Good

Bad

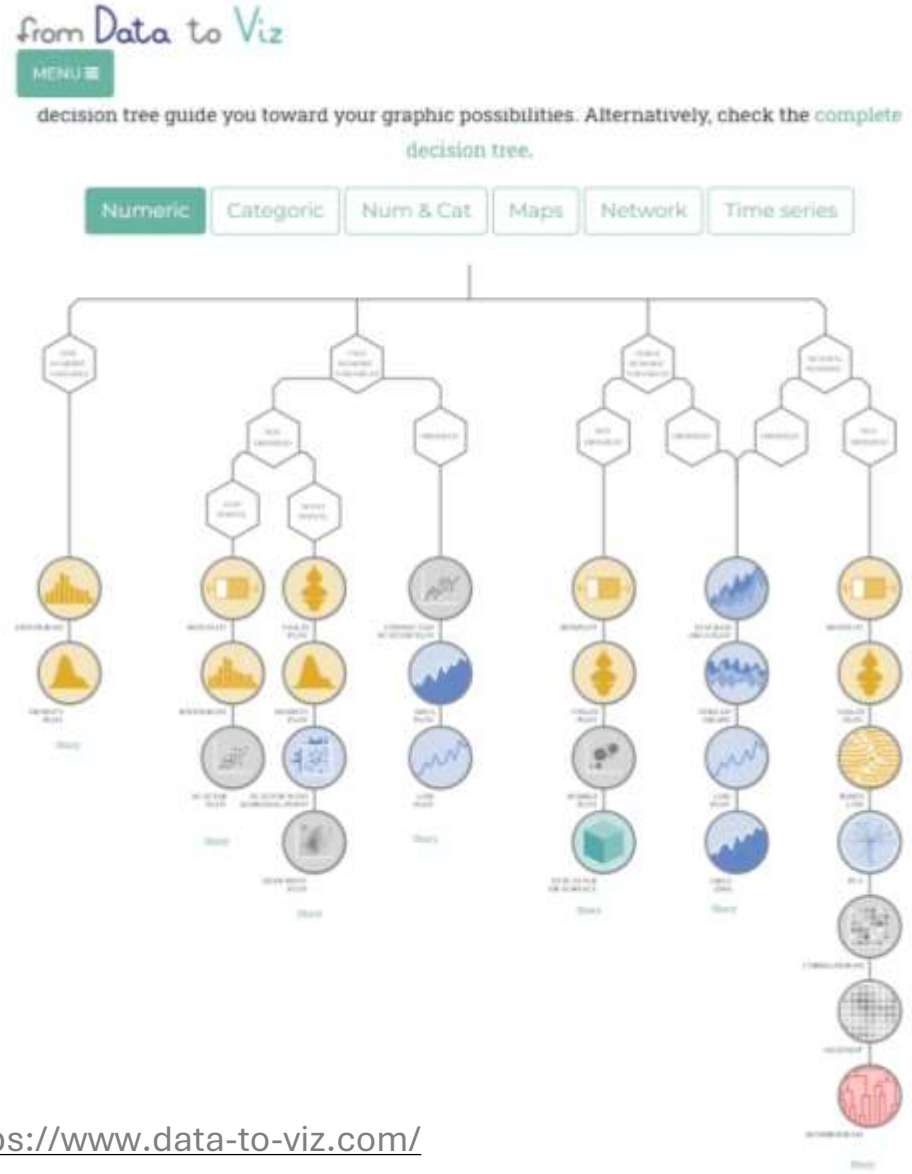
Banana

Sky

How: Choose the Right Chart Type



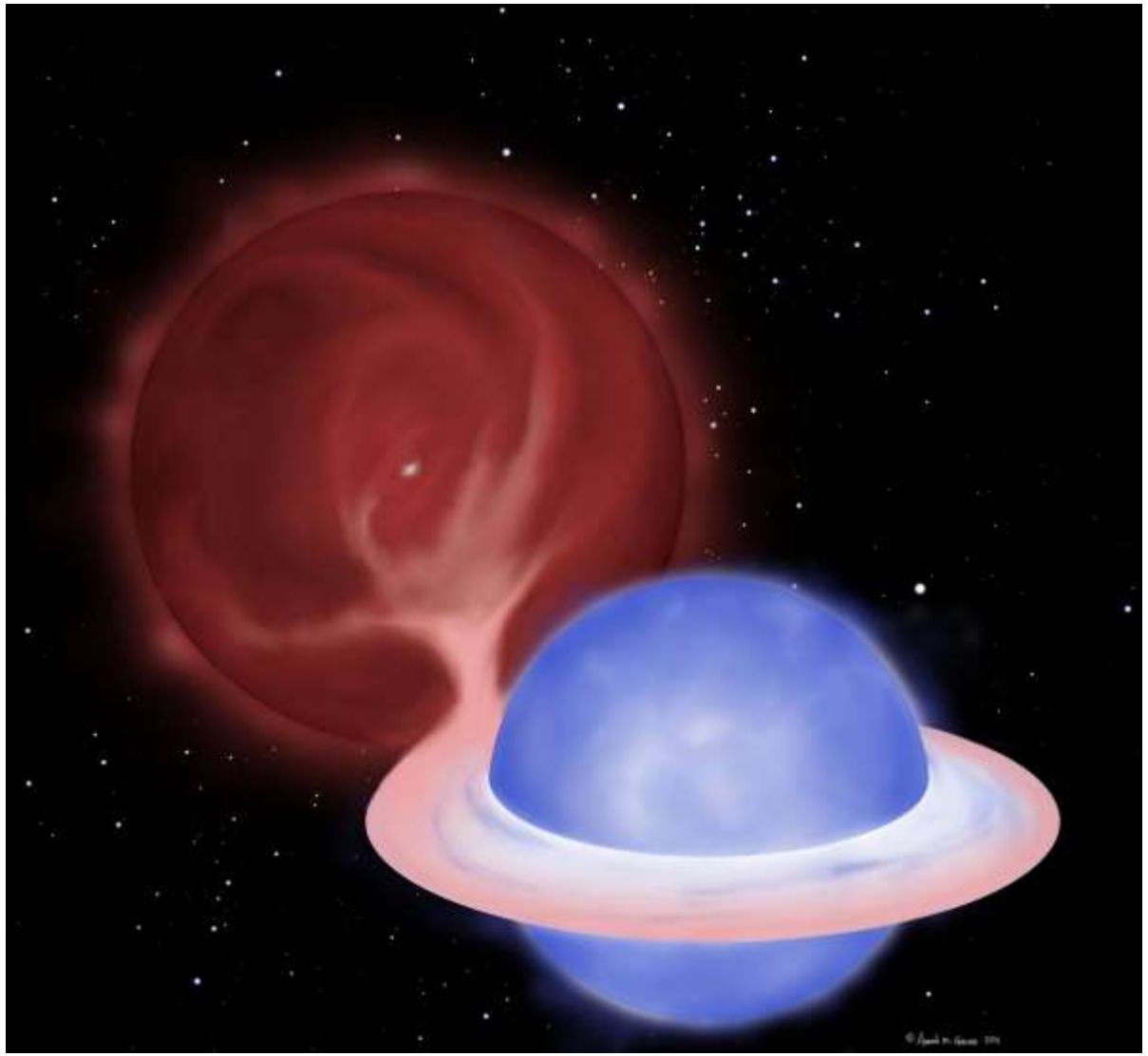
Abela (2006)



https://www.data-to-viz.com/

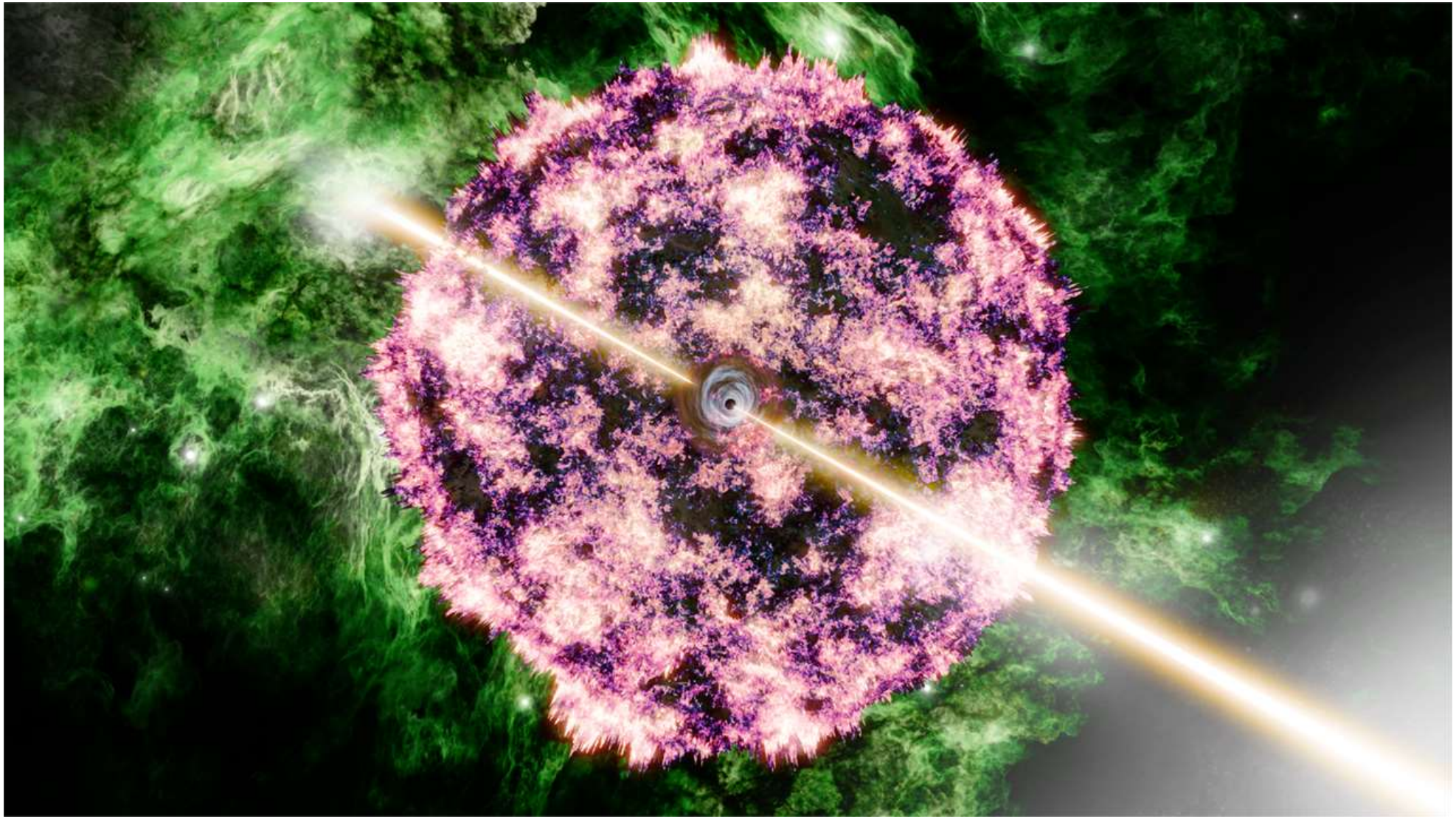
How: Artistic License

To facilitate science communication



How: Artistic License

For visual impact



Generative AI? …

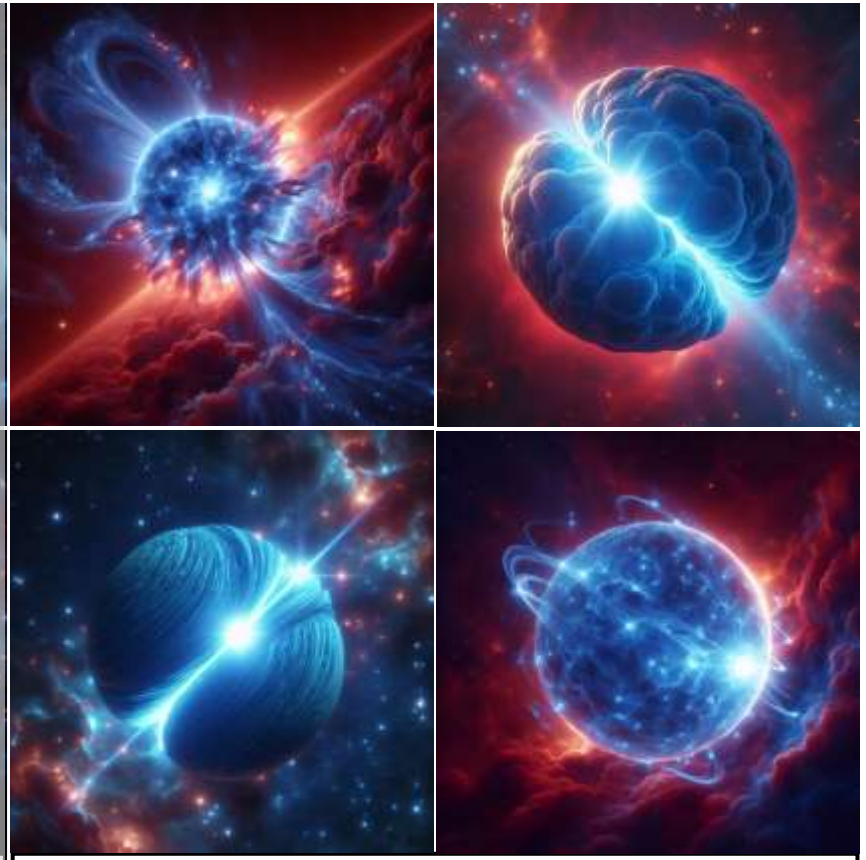
Images generated by DALL-E 3 via Microsoft Copilot

Prompt : “Please create an image of a cosmological volume with multiple galaxies with active star formation.”

Generative AI? …

Images generated by DALL-E 3 via Microsoft Copilot

Prompt : "Please create an image of a cosmological volume with multiple galaxies with active star formation.."



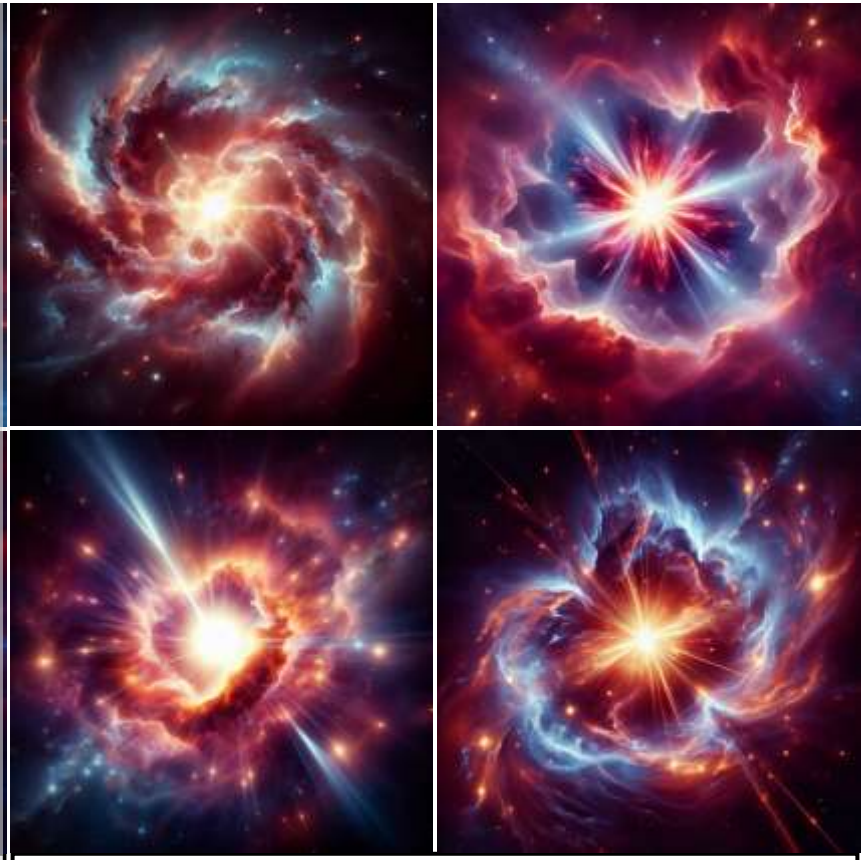
Prompt : "Please create an image of a blue straggler star forming by accreting mass from a red giant star."

Generative AI? …

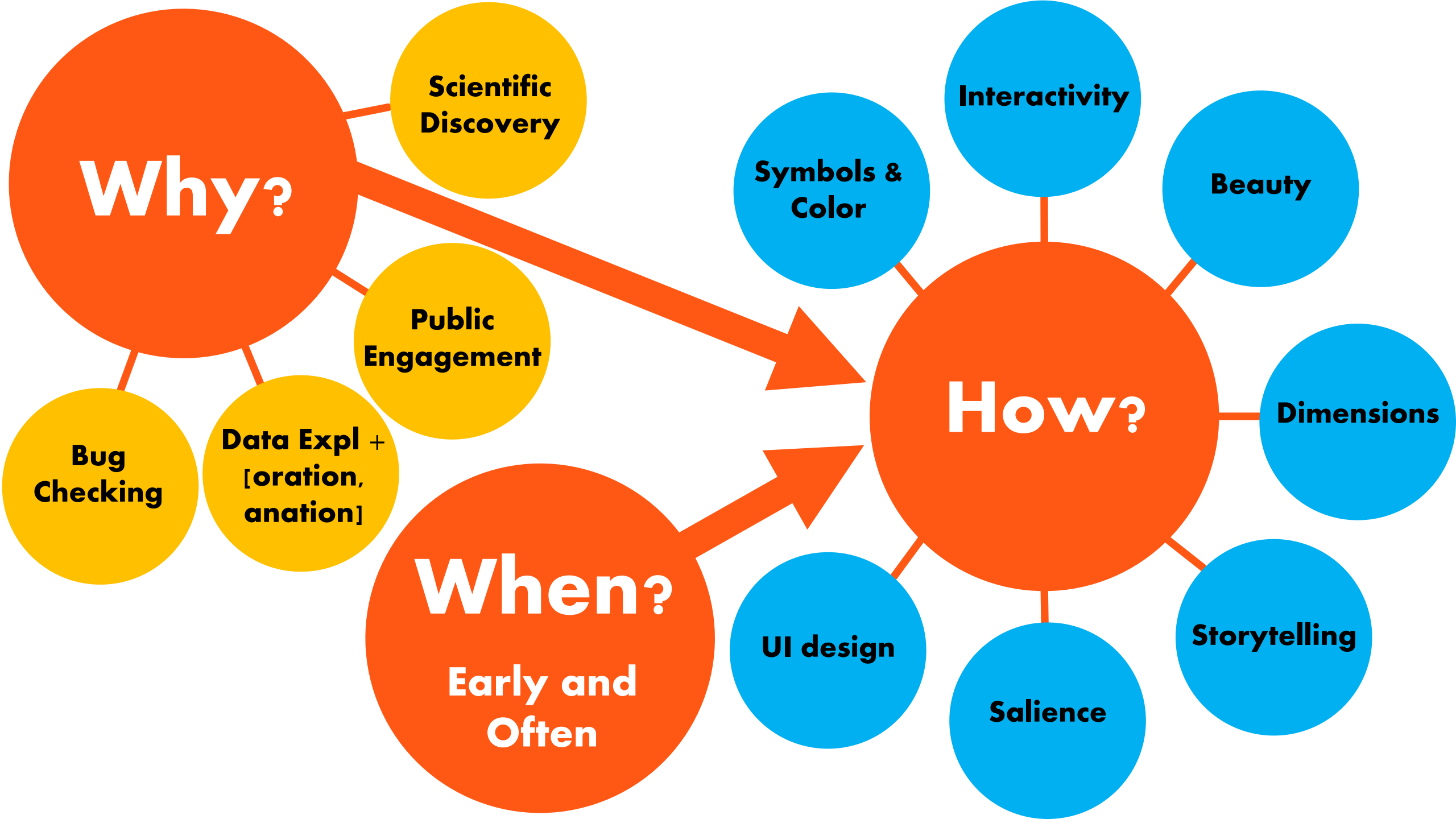
Images generated by DALL-E 3 via Microsoft Copilot

Prompt : "Please create an image of a cosmological volume with multiple galaxies with active star formation.."

Prompt : "Please create an image of a blue straggler star forming by accreting mass from a red giant star."



Prompt : "Please create an image of a supernova with a gamma ray burst inside a star forming region."


Why?
Scientific Discovery
Public Engagement
Bug Checking
Data Expl + [oration, anation]
When?
Early and Often
How?
Symbols & Color
Interactivity
Beauty
Dimensions
Storytelling
Salience
UI design

My Additional "Two Cents"

- **Interactivity** can enhance many of the "why's".

- We are in dire need of **innovative strategies for visualizing large data** (e.g., for Gaia, LSST, etc.)

Two great resources: Nature Methods points of view, Google material design principles

Questions?

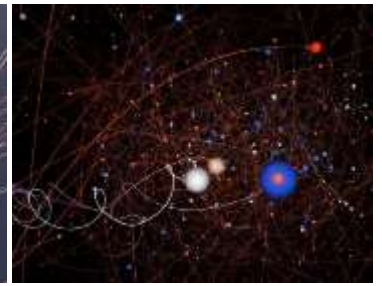
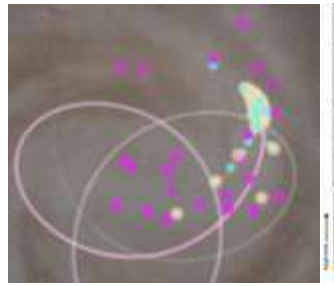
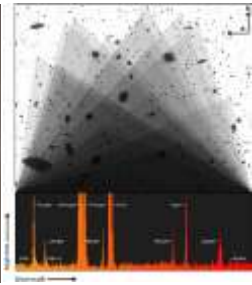
(I have a hands-on Python demo next.)

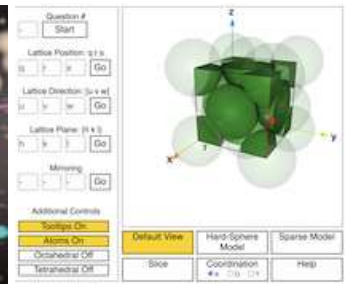
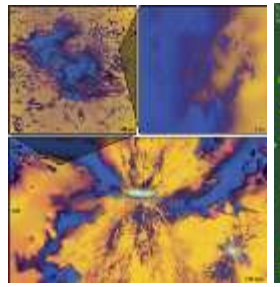
Exercise

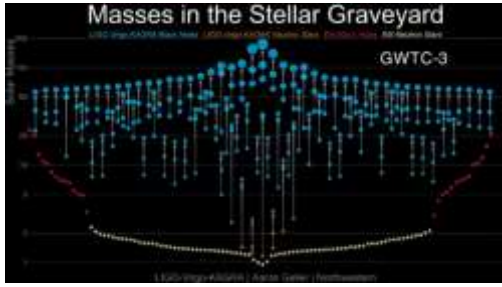
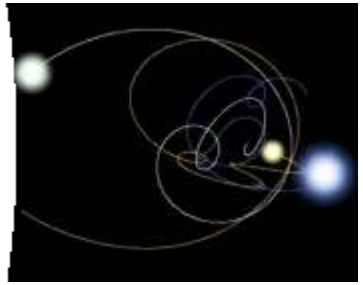
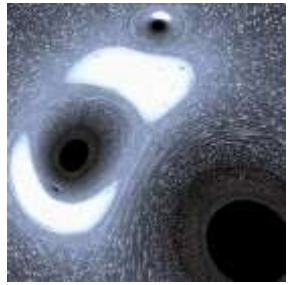
1. Either work with your own data, find a dataset online, or use a dataset from one of these resources:
   - Chicago Data Portal
   - Evanston Data Portal
   - A few datasets I have available on GitHub
2. Create a figure that tells a story.
   - Use the recommendations from this workshop!
   - Make a few drafts
3. Share your figure(s) with us and discuss the choices you made.

Thank you!

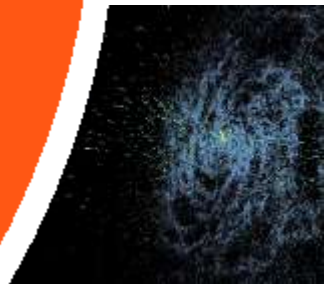
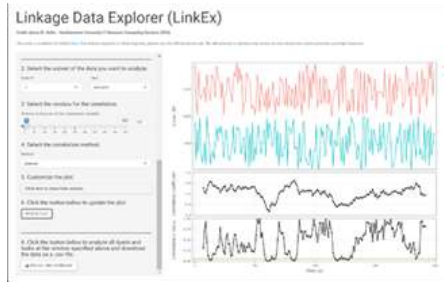


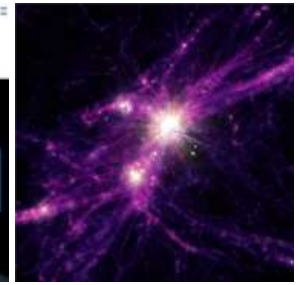





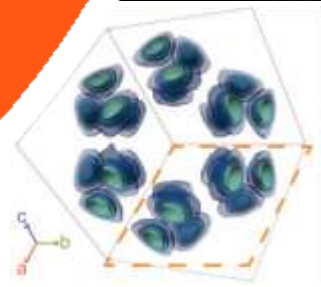
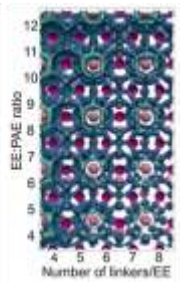

Start
Default View
Hard-Sphere Model
Sparse Model
Slice
Coordination
Help



Masses in the Stellar Graveyard
GWTC-3







Linkage Data Explorer (LinkEx)

CHiMaD
Center for Hierarchical Materials Design
CHiMaD System Design Toolbox




EE:PAE ratio
Number of linkers/EE