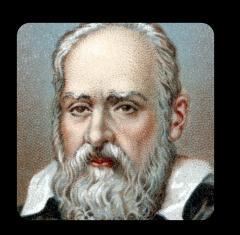


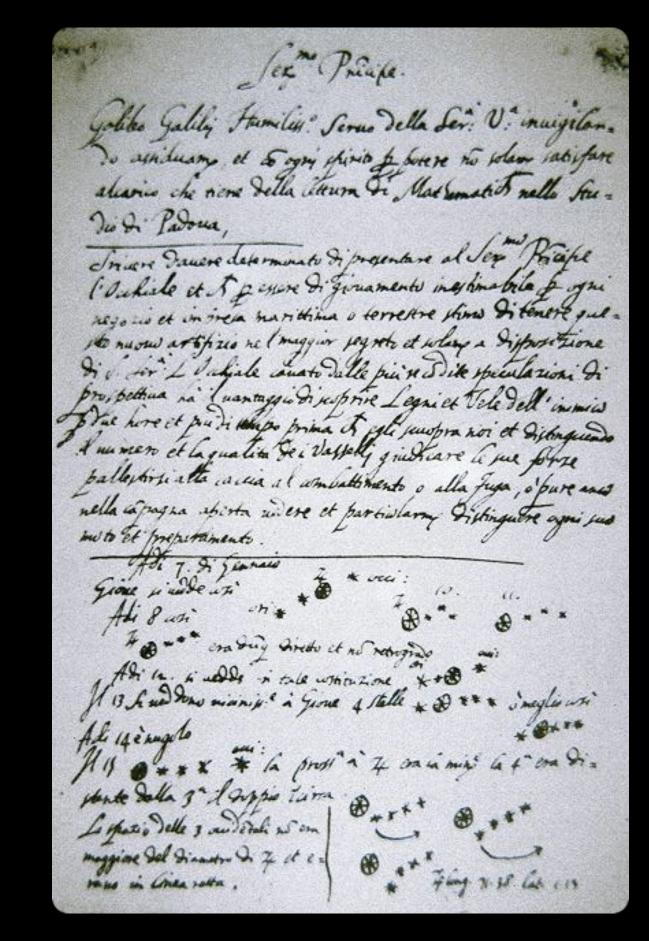
# glue-ing together the Universe from Galileo to Gaia



Alyssa A. Goodman

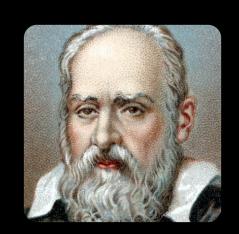


### Galileo, Jupiter's Moons, "3D" thinking



7 * *0 *	j # * O			
P 0 ***	* 0			
* * 0	9 . 0 .			
** •	4 * • • • •			
n + +O *	2# ·O· 'O. O.			
13. + 0.**	<b>1</b> 1O			
15 🔘 * * * *	(32			
15 O *	22 0.			
15 .0*	12)			
17 * 0 * 24	0 24 . 0			

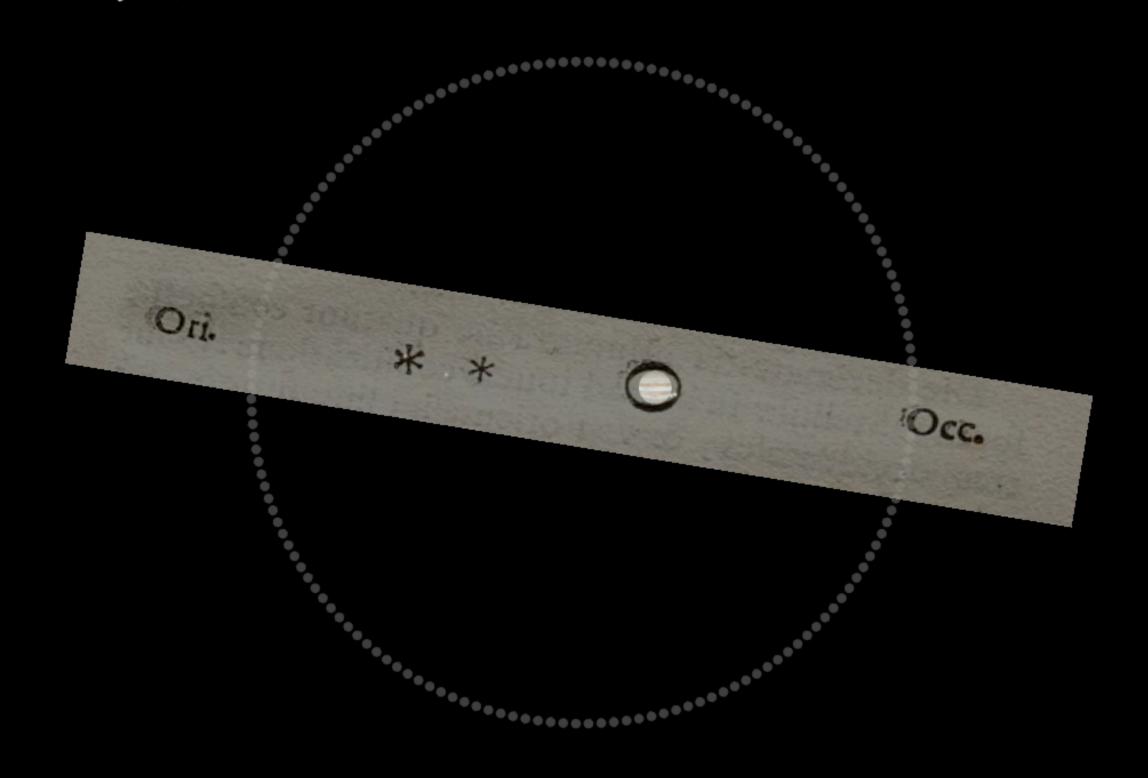
On the third sequence. The the closest wes		as I m	ninute, 30	seconds fr	om Jupiter;
East	*	× 0	*		* West
o minutes ren same straight li On the four Jupiter, two to	ine and of equ th, at the seco	nd ho	gnitude. ur, there v	vere four s	tars around
East	*	-0	*	*	West
and this one 6 m were nearly equ than the rest. I 30 seconds apa	But at the sever art. Jupiter wa	enth h	Jupiter ap our the ea ninutes fro	ppeared a l	ittle smaller s were only arer eastern
East	**	0	т.	*	West
one, while he	utes from the	wester	nmost on	e. They w	ere all equal
On the fifth	the sky was , tonly two sta	cloud	y.		
On the fifth	, the sky was	cloud	y.		



### Galileo's 3D thinking, in WorldWide Telescope

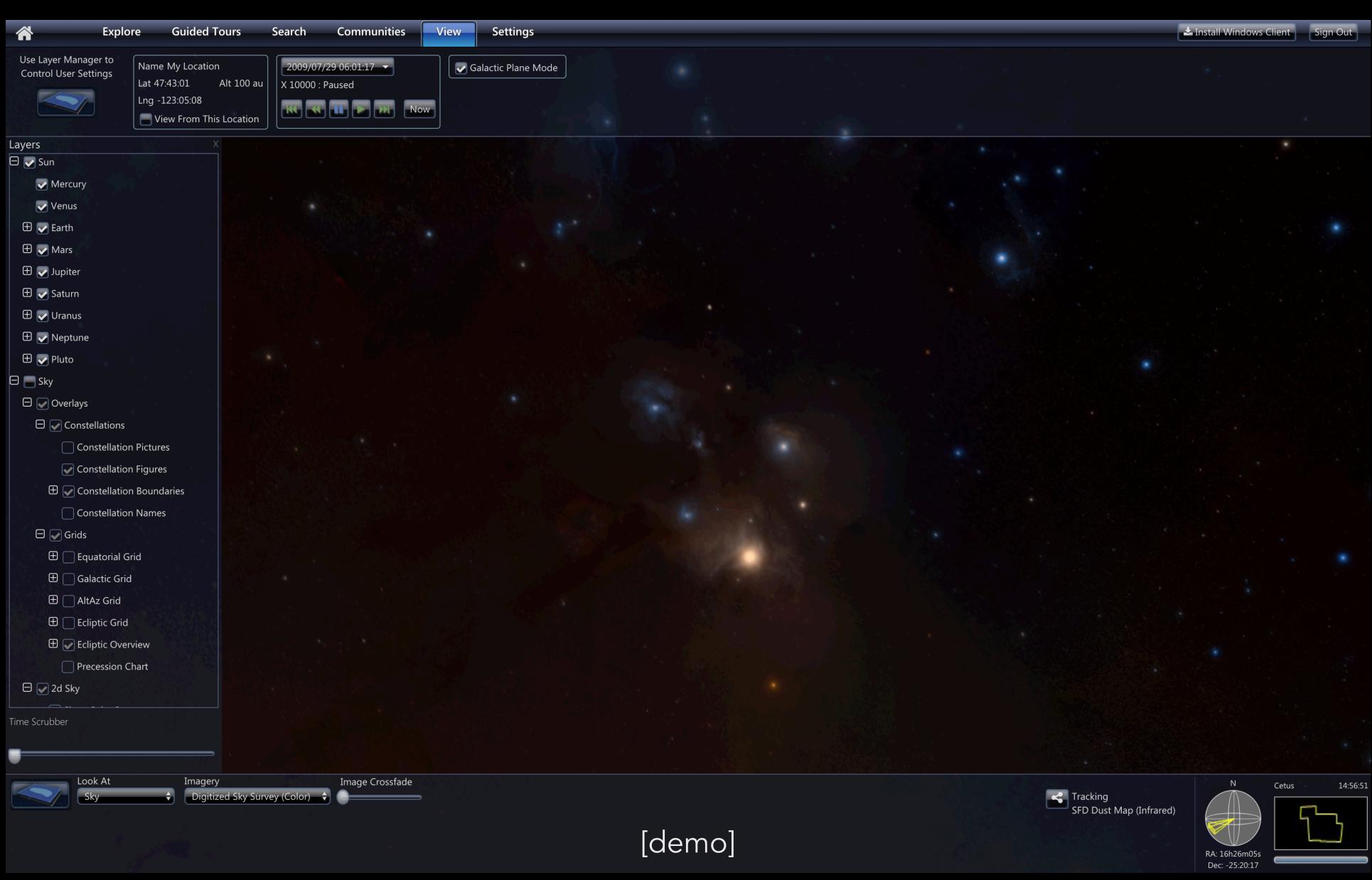


January 11, 1610

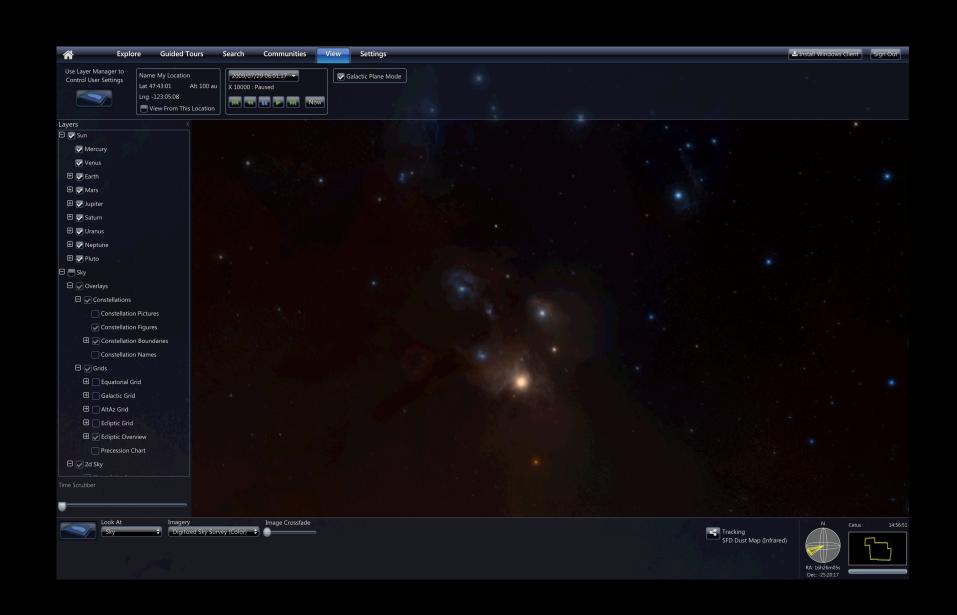


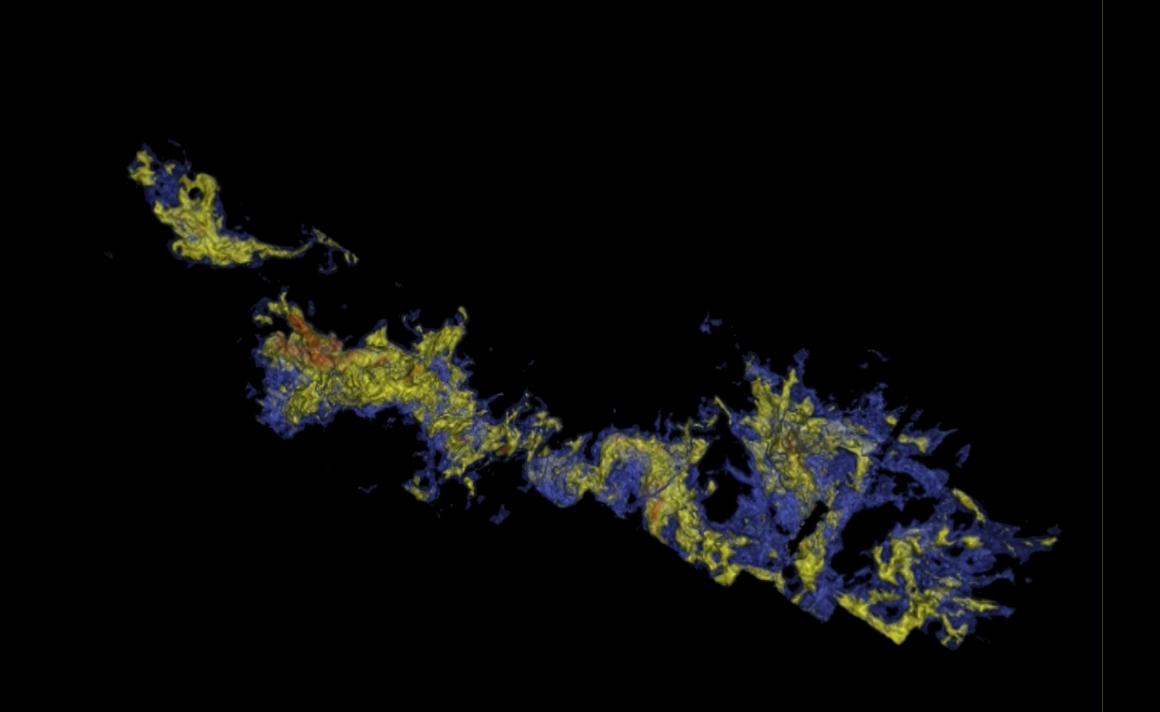
### The Sky at Many Wavelengths in a "WorldWide Telescope"

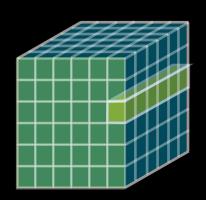


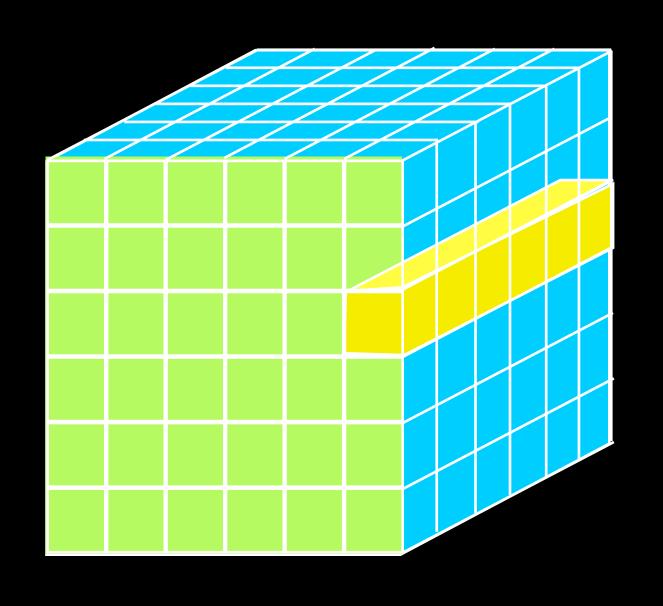


### Dimensions' many meanings









### "Data, Dimensions, Display"

1D: Columns = "Spectra", "SEDs" or "Time Series"

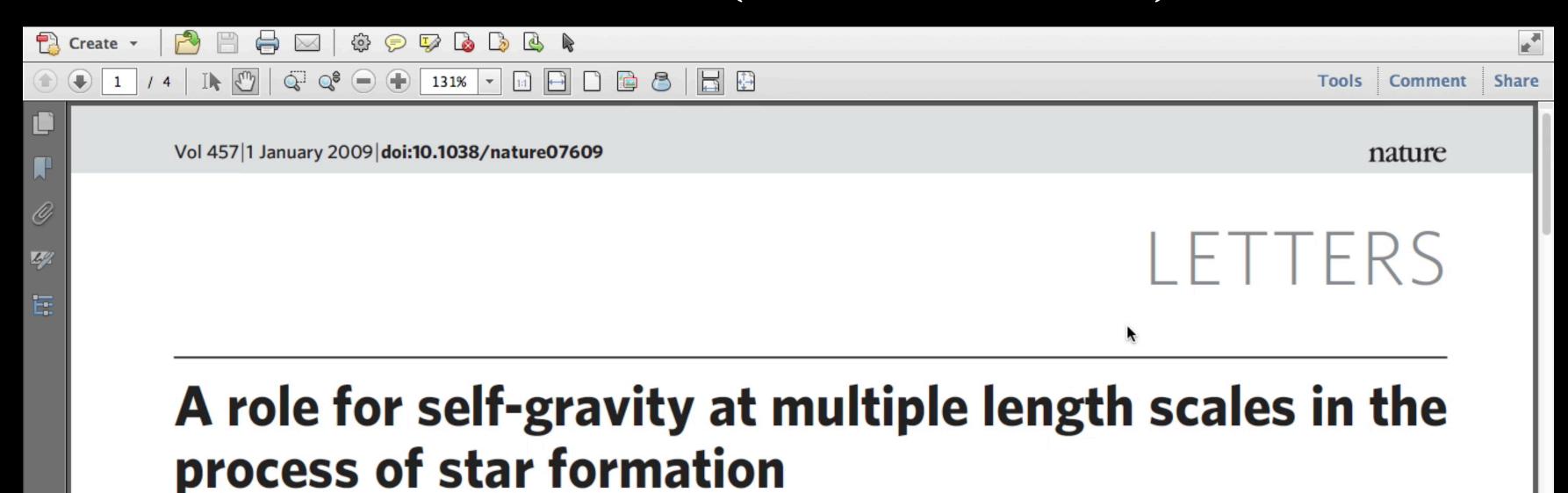
2D: Faces or Slices = "Images"

3D: Volumes = "3D Renderings", "2D Movies"

4D: Time Series of Volumes = "3D Movies"



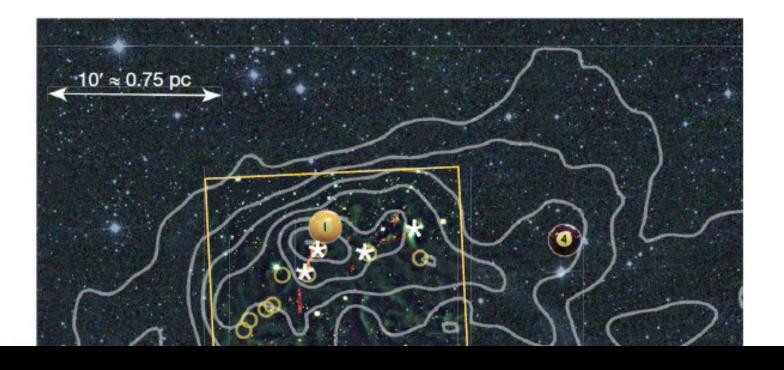
### "3D PDF" (Nature, 2009)

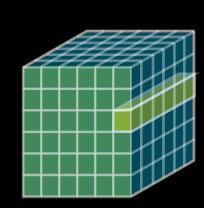


Alyssa A. Goodman<sup>1,2</sup>, Erik W. Rosolowsky<sup>2,3</sup>, Michelle A. Borkin<sup>1</sup>†, Jonathan B. Foster<sup>2</sup>, Michael Halle<sup>1,4</sup>, Jens Kauffmann<sup>1,2</sup> & Jaime E. Pineda<sup>2</sup>

Self-gravity plays a decisive role in the final stages of star formation, where dense cores (size ~0.1 parsecs) inside molecular clouds collapse to form star-plus-disk systems1. But self-gravity's role at earlier times (and on larger length scales, such as ~1 parsec) is unclear; some molecular cloud simulations that do not include self-gravity suggest that 'turbulent fragmentation' alone is sufficient to create a mass distribution of dense cores that resembles, and sets, the stellar initial mass function2. Here we report a 'dendrogram' (hierarchical tree-diagram) analysis that reveals that self-gravity plays a significant role over the full range of possible scales traced by <sup>13</sup>CO observations in the L1448 molecular cloud, but not everywhere in the observed region. In particular, more than 90 per cent of the compact 'pre-stellar cores' traced by peaks of dust emission3 are projected on the sky within one of the dendrogram's self-gravitating 'leaves'. As these peaks mark the locations of already-forming stars, or of those probably about to form, a self-gravitating cocoon seems a critical condition for their exist.

overlapping features as an option, significant emission found between prominent clumps is typically either appended to the nearest clump or turned into a small, usually 'pathological', feature needed to encompass all the emission being modelled. When applied to molecular-line





### 2009 3D PDF

LETTERS NATURE|Vol 457|1 January 2009

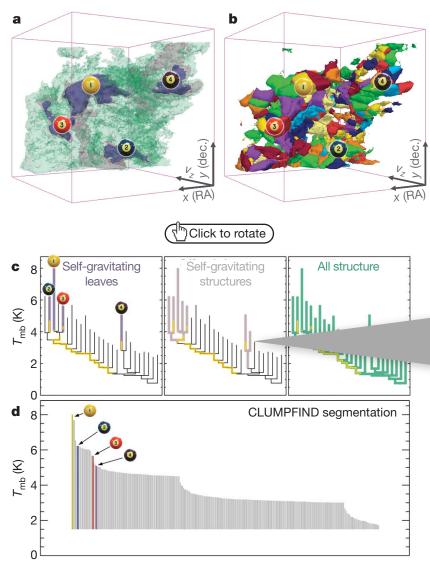


Figure 2 | Comparison of the 'dendrogram' and 'CLUMPFIND' featureidentification algorithms as applied to <sup>13</sup>CO emission from the L1448 region of Perseus. a, 3D visualization of the surfaces indicated by colours in the dendrogram shown in c. Purple illustrates the smallest scale selfgravitating structures in the region corresponding to the leaves of the dendrogram; pink shows the smallest surfaces that contain distinct selfgravitating leaves within them; and green corresponds to the surface in the data cube containing all the significant emission. Dendrogram branches corresponding to self-gravitating objects have been highlighted in yellow over the range of  $T_{\rm mb}$  (main-beam temperature) test-level values for which the virial parameter is less than 2. The x-y locations of the four 'selfgravitating' leaves labelled with billiard balls are the same as those shown in Fig. 1. The 3D visualizations show position–position–velocity  $(p-p-\nu)$  space. RA, right ascension; dec., declination. For comparison with the ability of dendrograms (c) to track hierarchical structure, d shows a pseudodendrogram of the CLUMPFIND segmentation (b), with the same four labels used in Fig. 1 and in a. As 'clumps' are not allowed to belong to larger structures, each pseudo-branch in **d** is simply a series of lines connecting the maximum emission value in each clump to the threshold value. A very large number of clumps appears in **b** because of the sensitivity of CLUMPFIND to noise and small-scale structure in the data. In the online PDF version, the 3D cubes (a and b) can be rotated to any orientation, and surfaces can be turned on and off (interaction requires Adobe Acrobat version 7.0.8 or higher). In the printed version, the front face of each 3D cube (the 'home' view in the interactive online version) corresponds exactly to the patch of sky shown in Fig. 1, and velocity with respect to the Local Standard of Rest increases from front  $(-0.5 \,\mathrm{km \, s}^{-1})$  to back  $(8 \,\mathrm{km \, s}^{-1})$ .

data, CLUMPFIND typically finds features on a limited range of scales, above but close to the physical resolution of the data, and its results can be overly dependent on input parameters. By tuning CLUMPFIND's two free parameters, the same molecular-line data set<sup>8</sup> can be used to show either that the frequency distribution of clump mass is the same as the initial mass function of stars or that it follows the much shallower mass function associated with large-scale molecular clouds (Supplementary Fig. 1).

Four years before the advent of CLUMPFIND, 'structure trees' were proposed as a way to characterize clouds' hierarchical structure

using 2D maps of column density. With the tion, we have developed a structure-id abstracts the hierarchical structure of a an easily visualized representation called well developed in other data-intensive application of tree methodologies so fa and almost exclusively within the atomic merger trees' are being used with in Figure 3 and its legend explain the

reger trees' are being used with in Figure 3 and its legend explain the paratically. The dendrogram and "Crad" panels!

That's not

Cand en and the second e

arly 2D work as inspira-

mass, suc. K<sup>-1</sup> km<sup>-1</sup> s

ary Fig. 2).

used to estimate the role of self-gravity at each point in the hierarchy, via calculation of an 'observed' virial parameter,  $\alpha_{\rm obs} = 5\sigma_{\nu}^{\ 2}R/GM_{\rm lum}$ . In principle, extended portions of the tree (Fig. 2, yellow highlighting) where  $\alpha_{\rm obs} < 2$  (where gravitational energy is comparable to or larger than kinetic energy) correspond to regions of  $p-p-\nu$  space where self-gravity is significant. As  $\alpha_{\rm obs}$  only represents the ratio of kinetic energy to gravitational energy at one point in time, and does not explicitly capture external over-pressure and/or magnetic fields<sup>16</sup>, its measured value should only be used as a guide to the longevity (boundedness) of any particular feature.

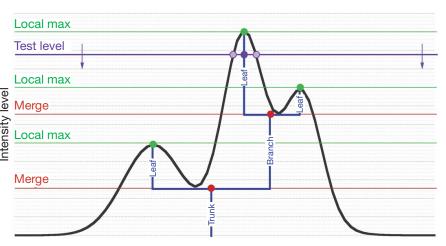
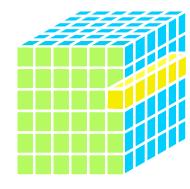
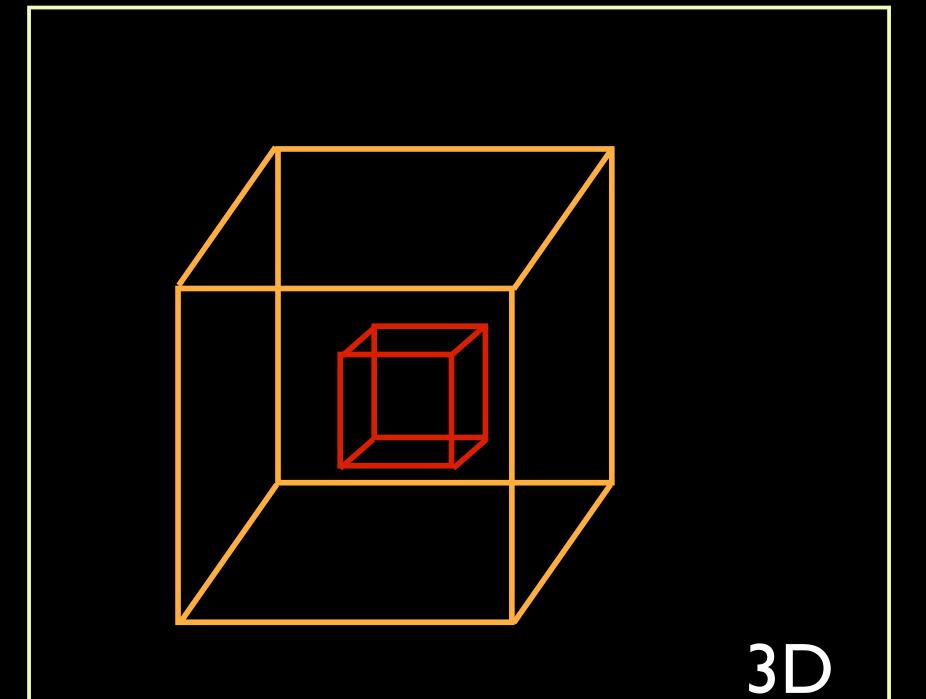


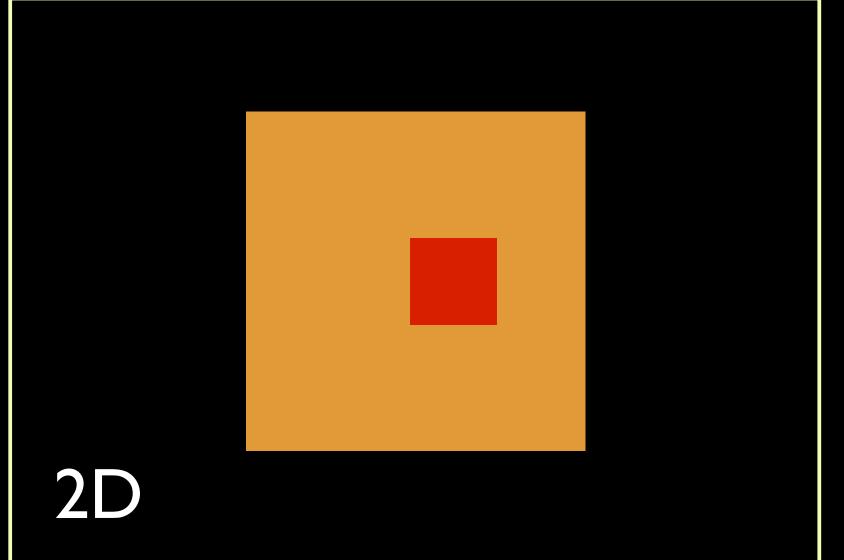
Figure 3 | Schematic illustration of the dendrogram process. Shown is the construction of a dendrogram from a hypothetical one-dimensional emission profile (black). The dendrogram (blue) can be constructed by 'dropping' a test constant emission level (purple) from above in tiny steps (exaggerated in size here, light lines) until all the local maxima and mergers are found, and connected as shown. The intersection of a test level with the emission is a set of points (for example the light purple dots) in one dimension, a planar curve in two dimensions, and an isosurface in three dimensions. The dendrogram of 3D data shown in Fig. 2c is the direct analogue of the tree shown here, only constructed from 'isosurface' rather than 'point' intersections. It has been sorted and flattened for representation on a flat page, as fully representing dendrograms for 3D data cubes would require four dimensions.



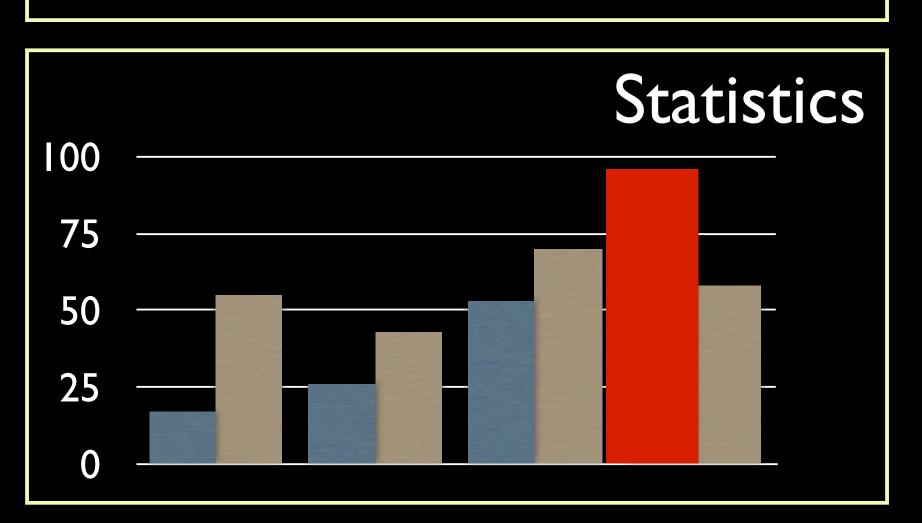
Goodman et al. 2009, Nature, cf: Fluke et al. 2009

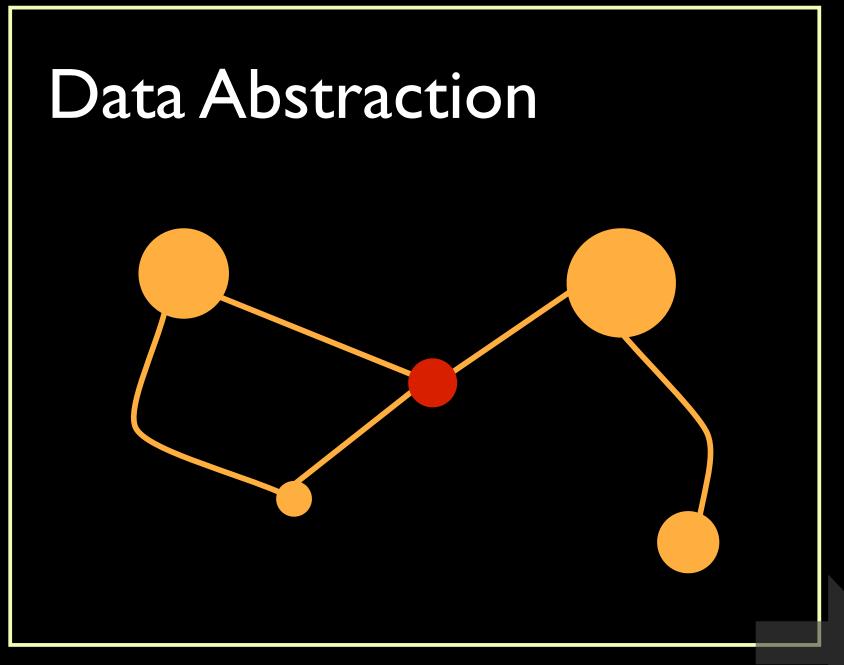
### Linked Views of High-dimensional Data





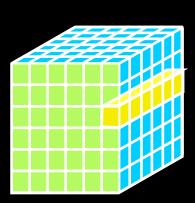


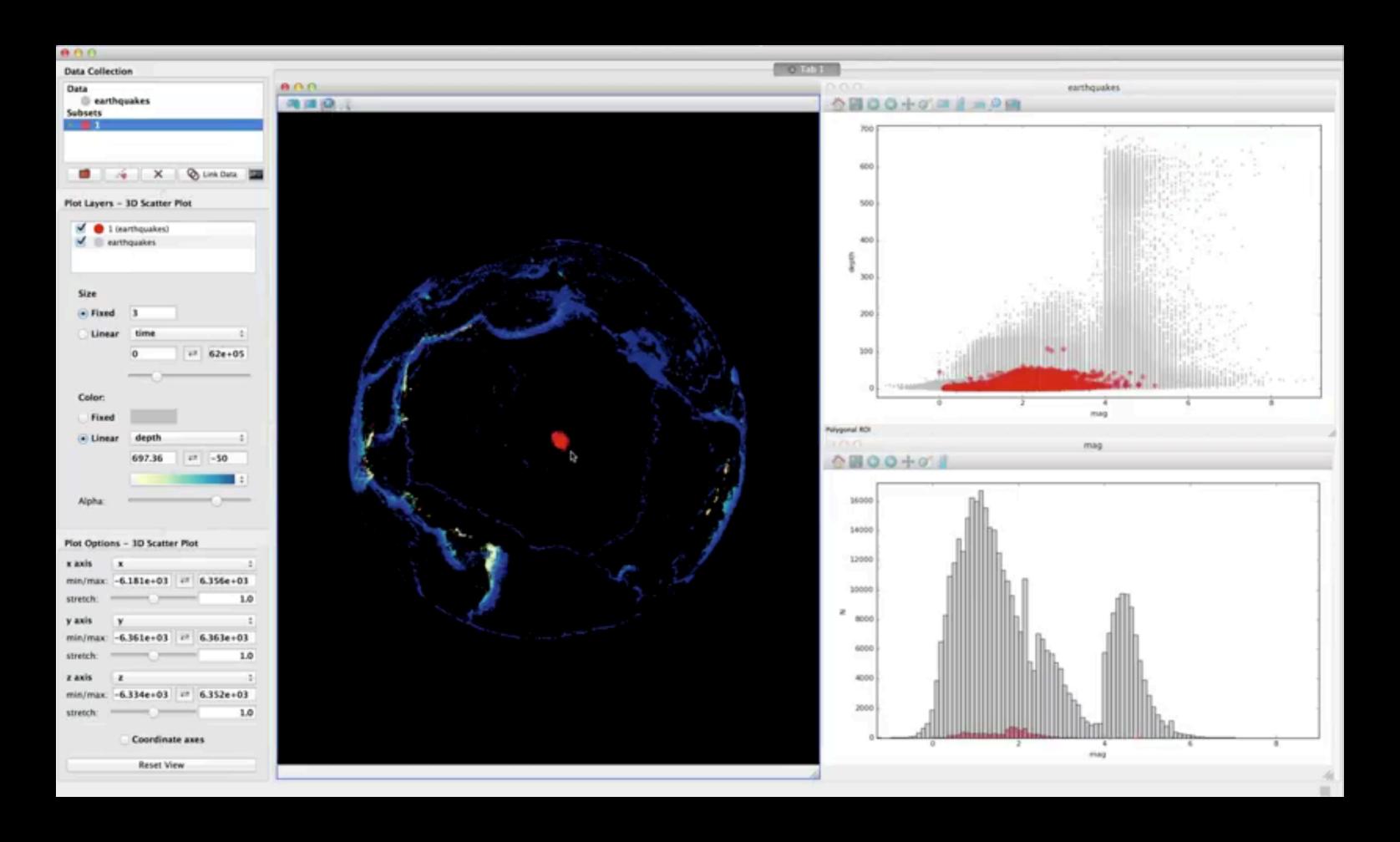


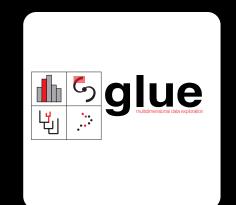


### Linked Views of High-dimensional Data (in Python)

### glue

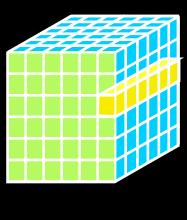


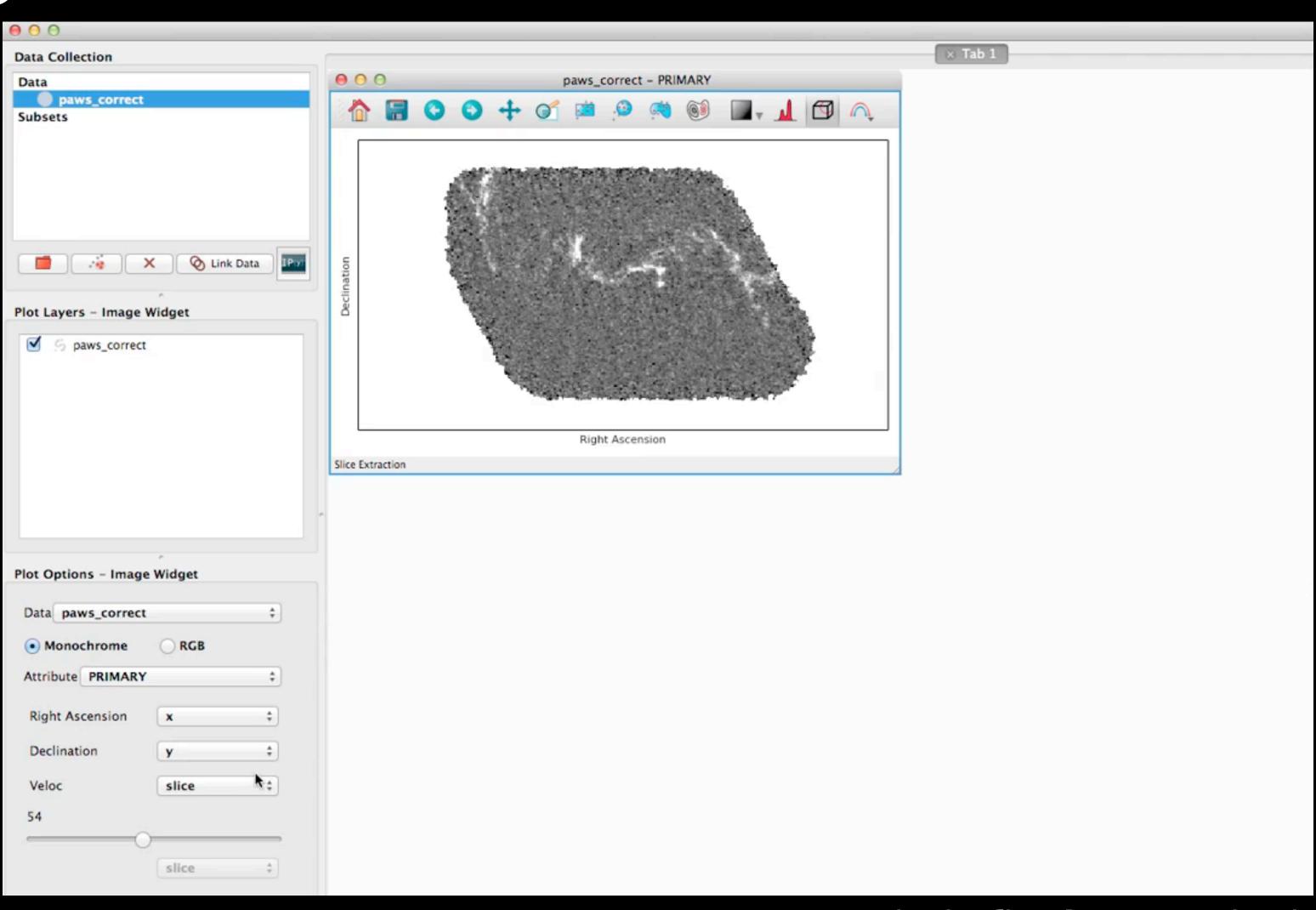


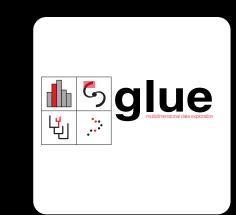


### Linked Views of High-dimensional Data (in Python)

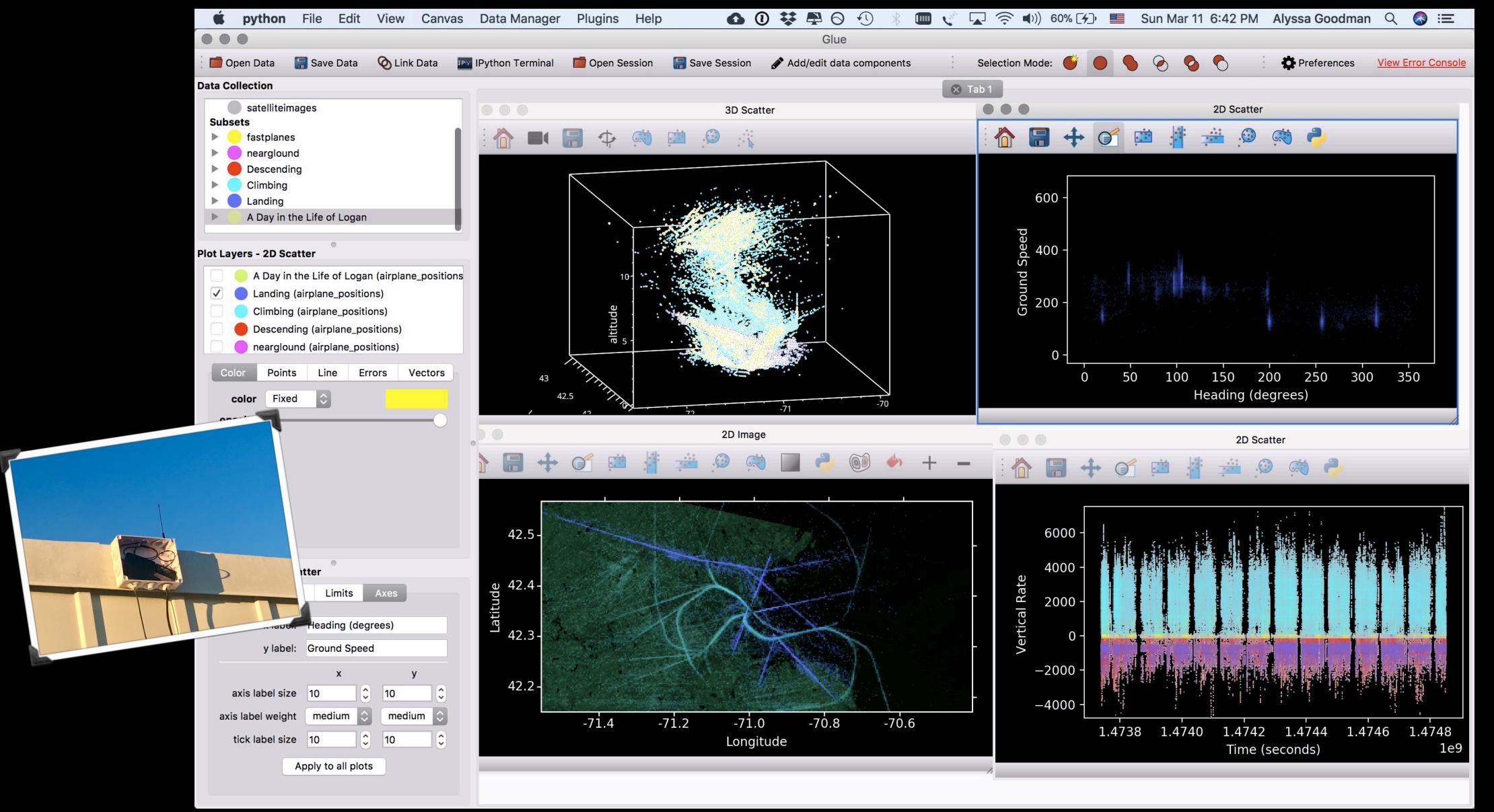
### glue

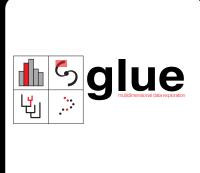


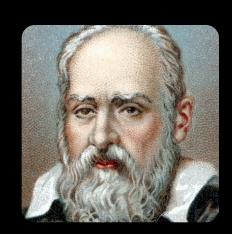




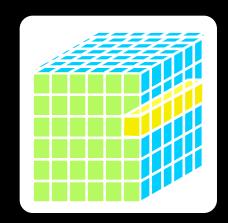
### Linked views & my FBI file...(glue's not just for Astronomy)



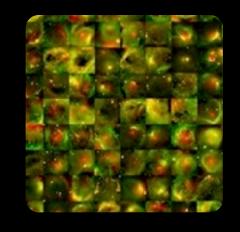


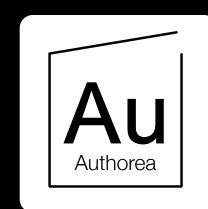


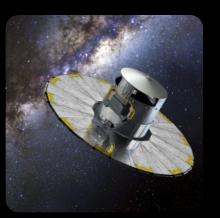




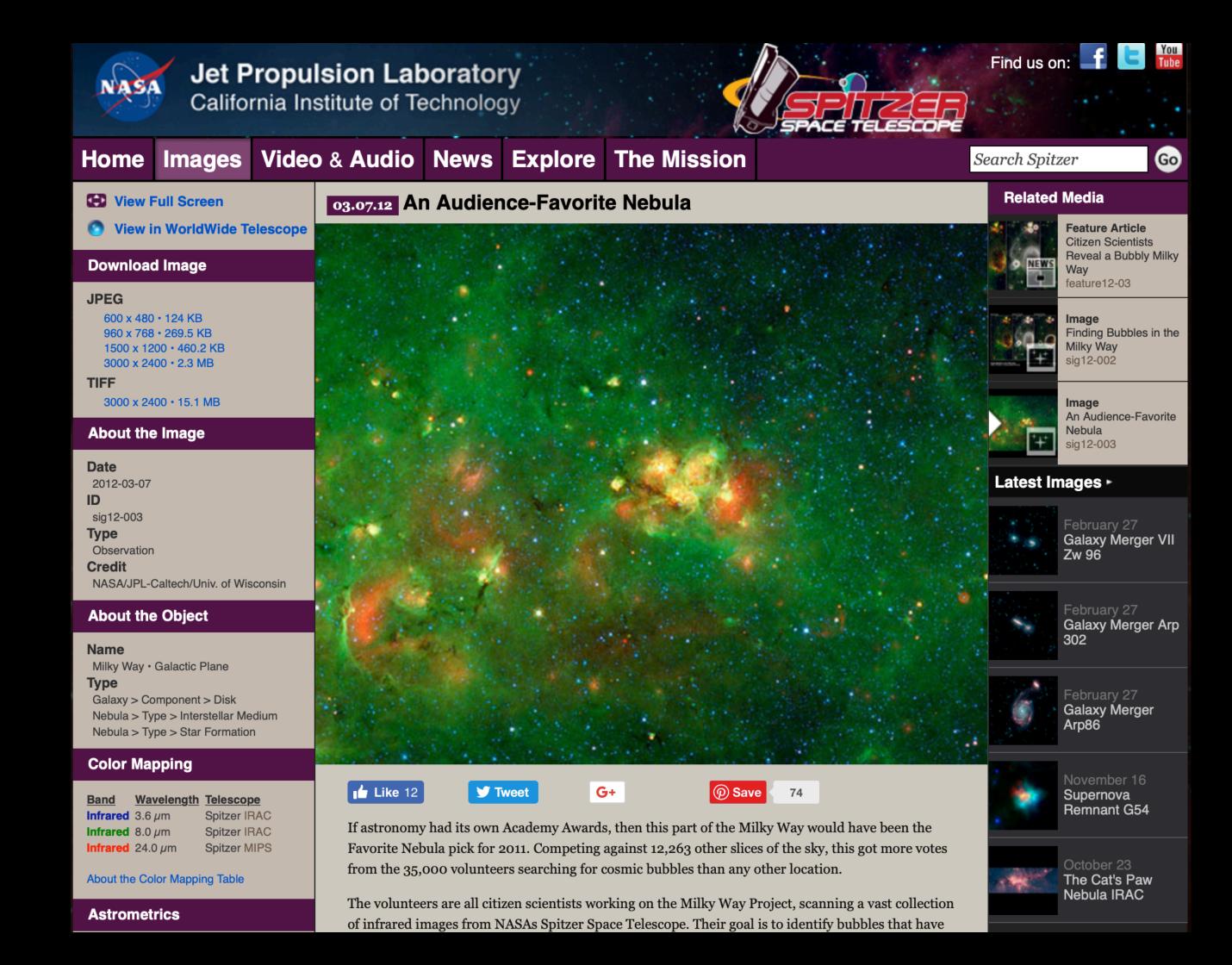


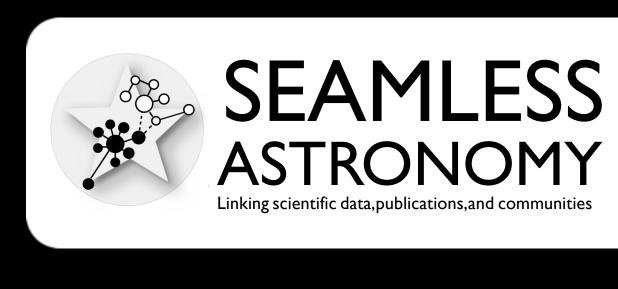


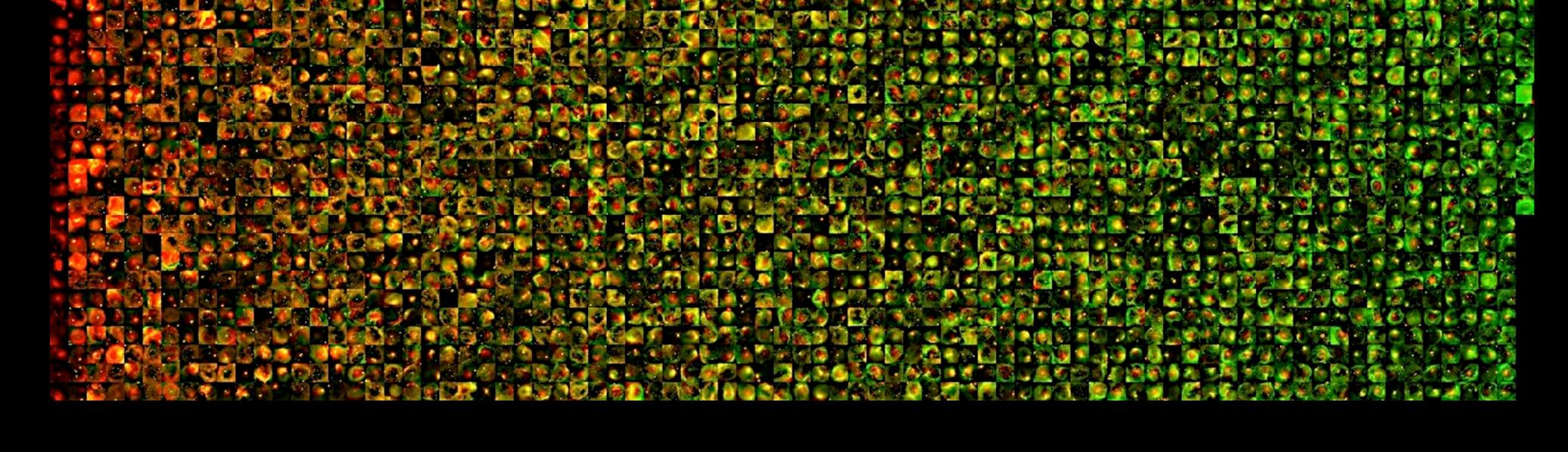




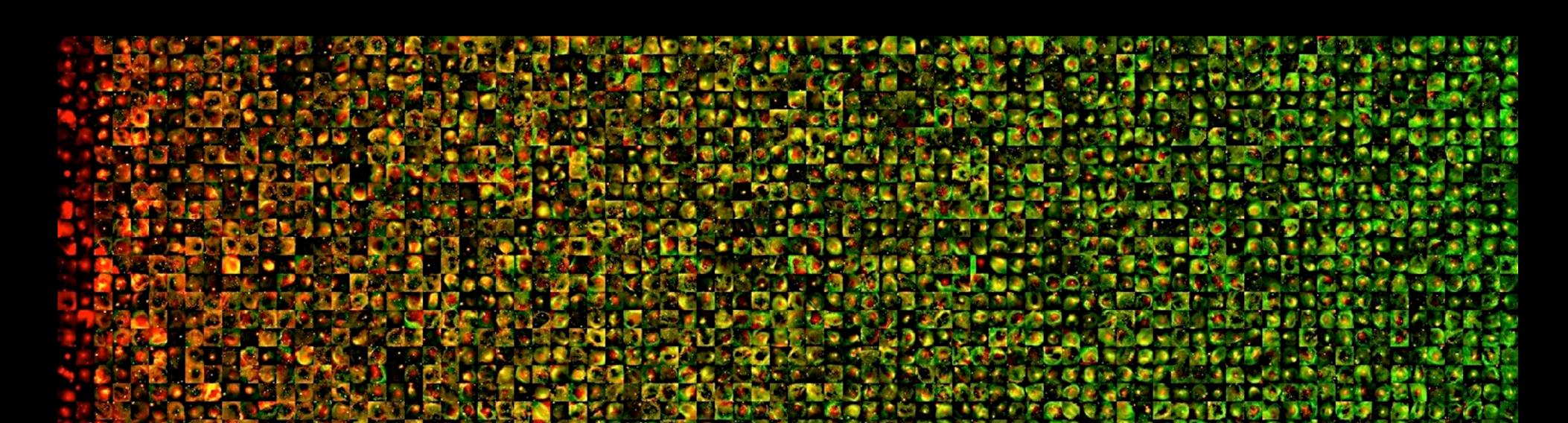


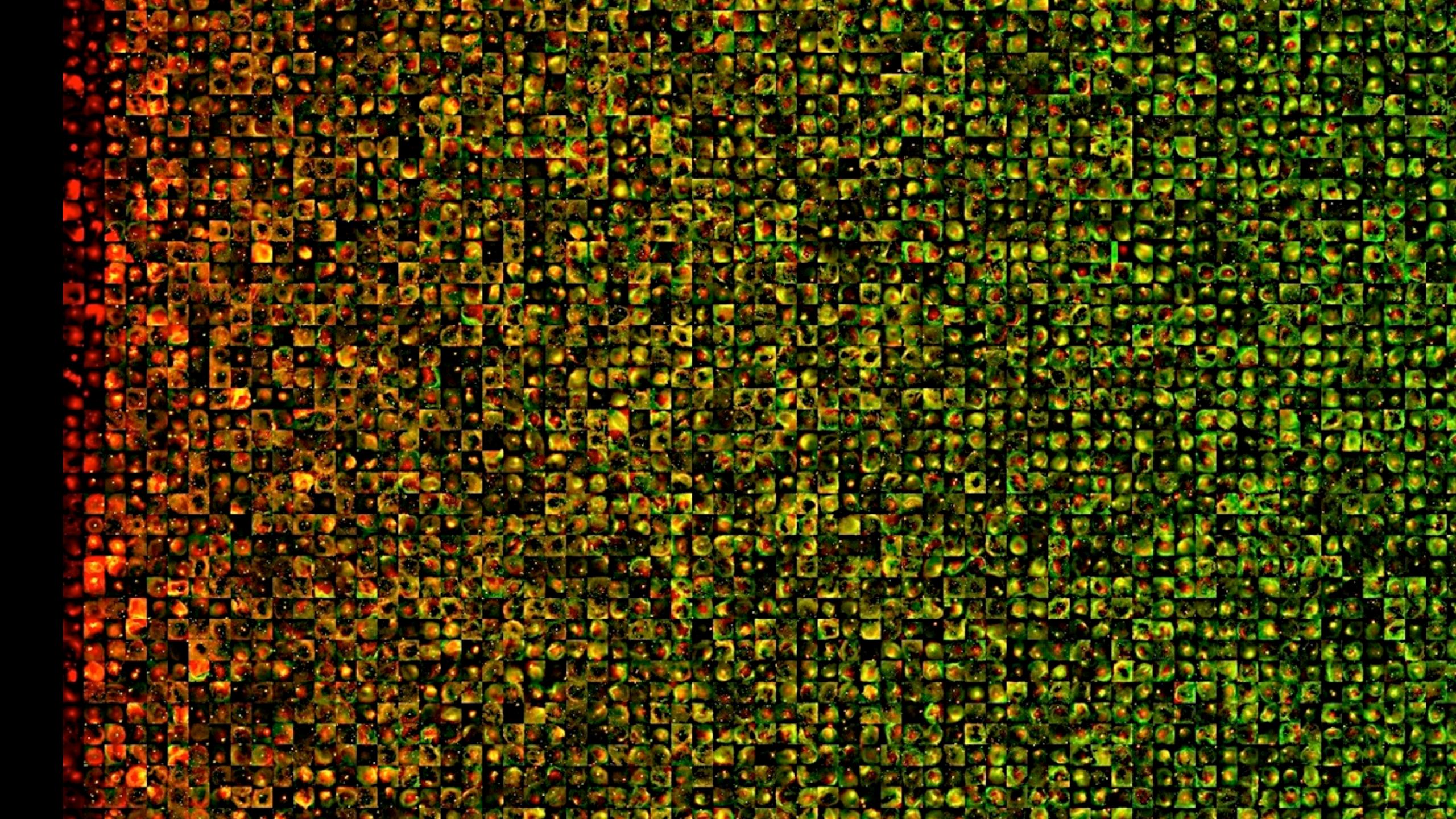




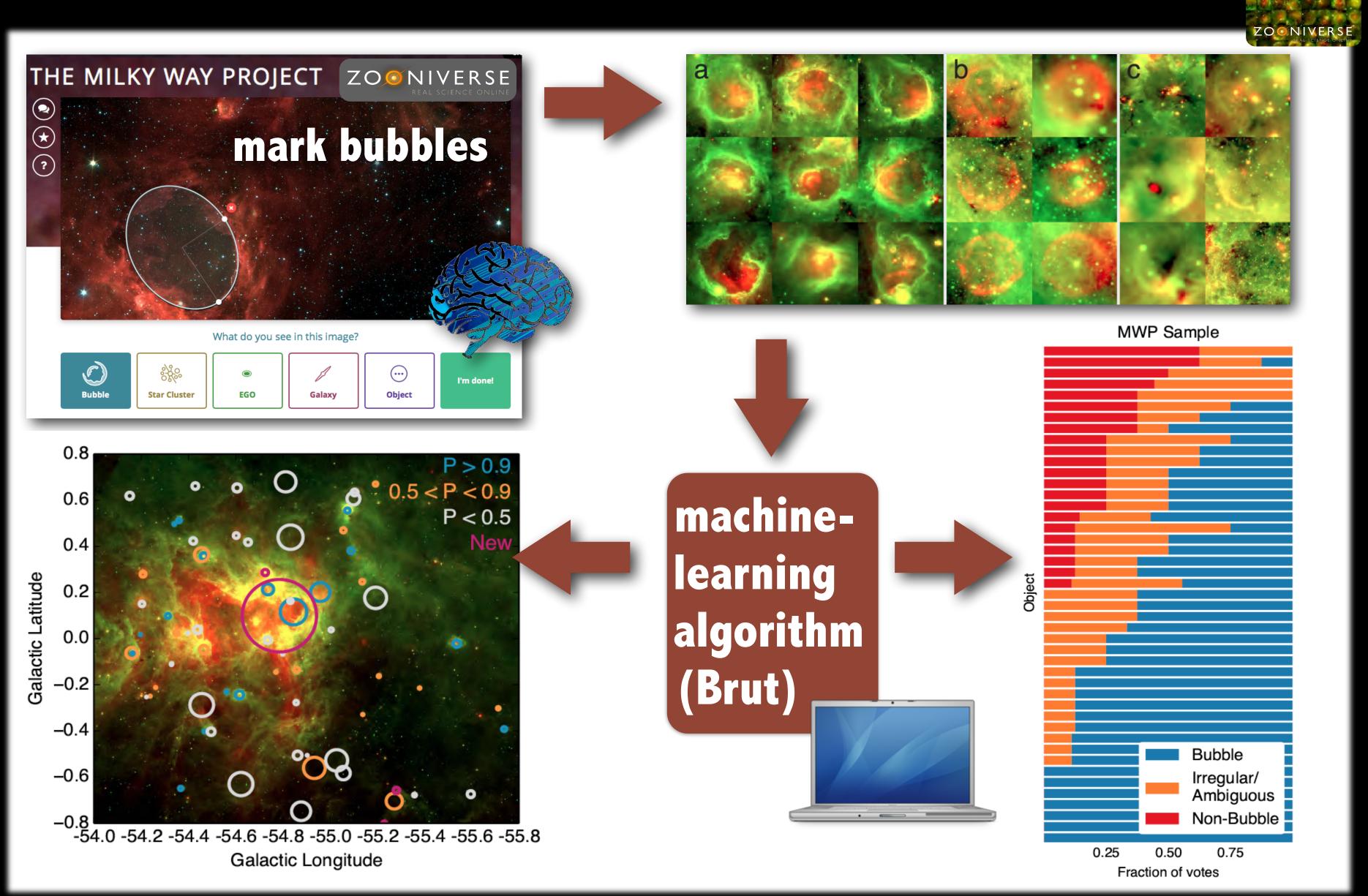


# BIG DATA, WIDE DATA

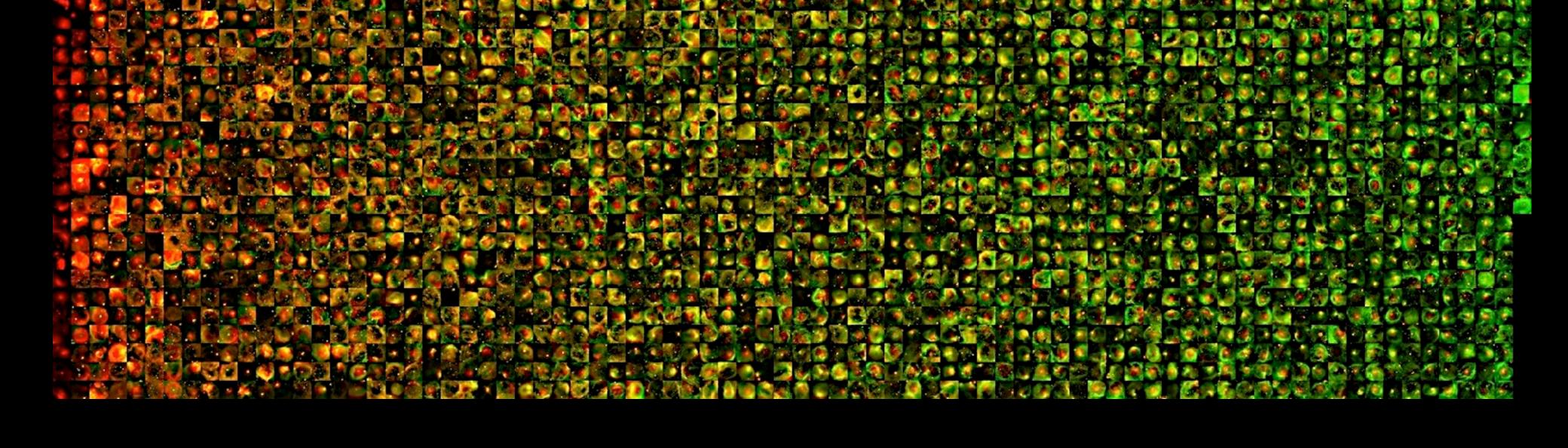




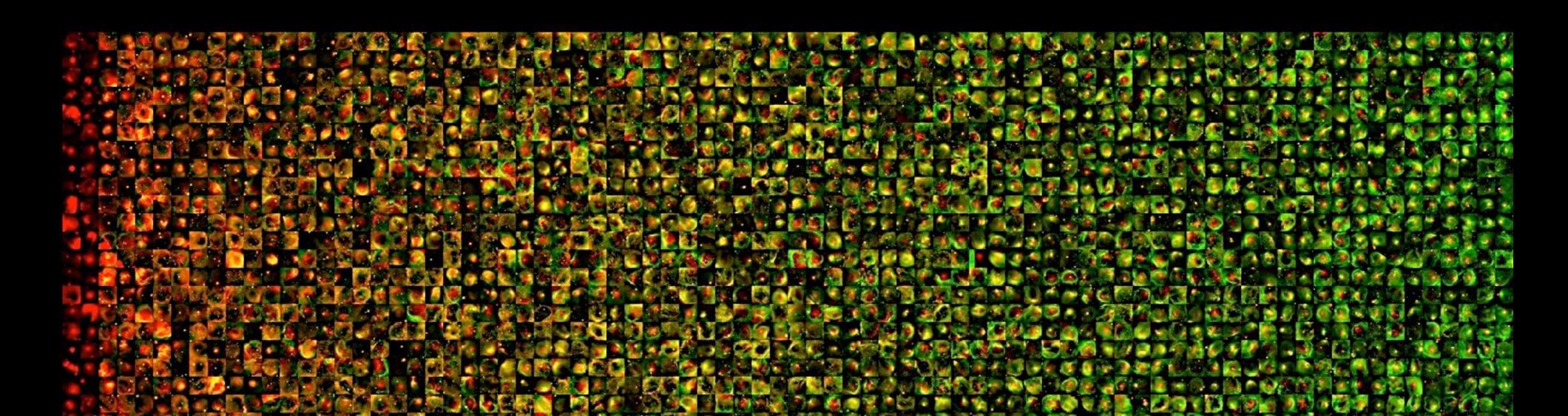
### BIG DATA AND "HUMAN-AIDED COMPUTING"



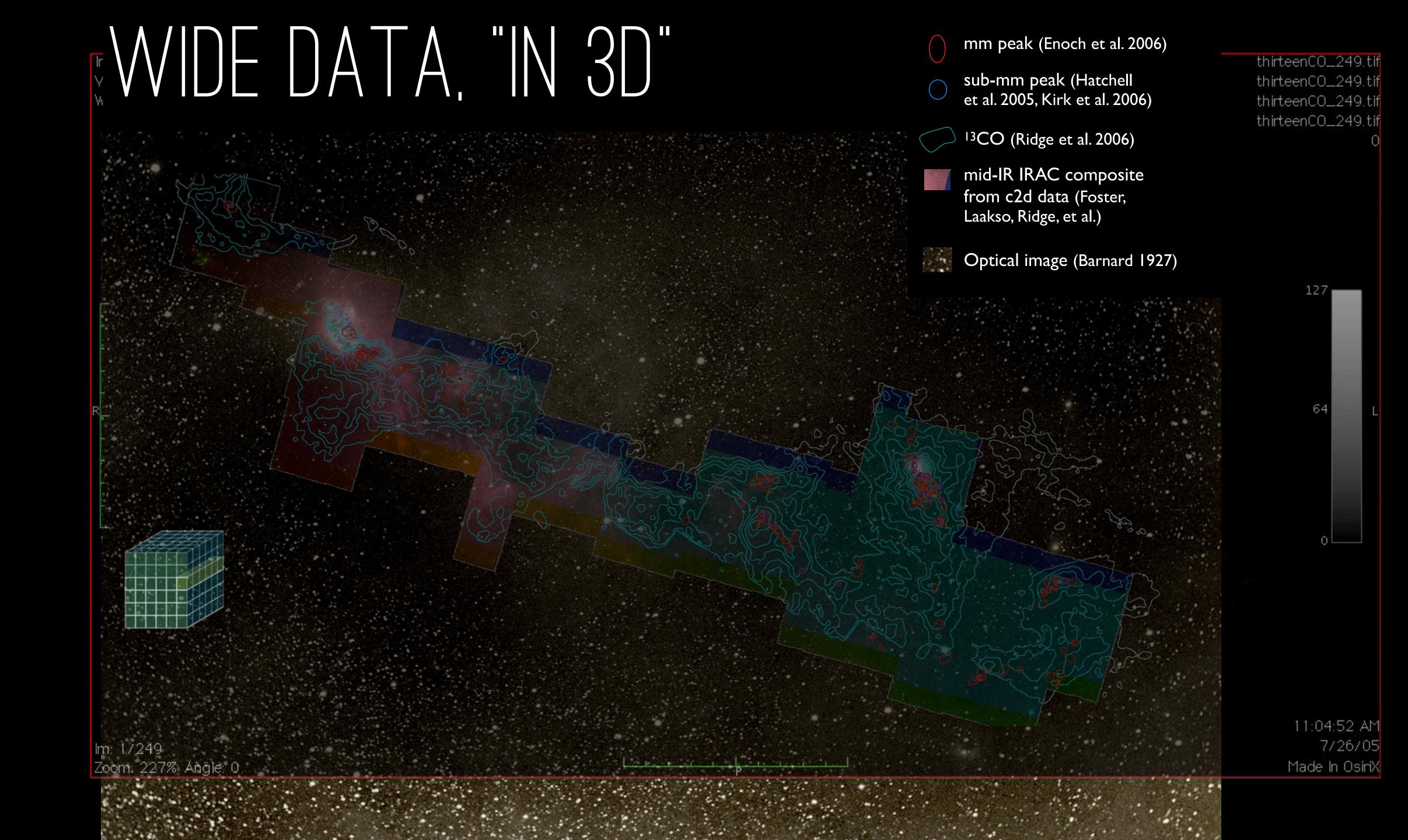
example here from: **Beaumont**, Goodman, Kendrew, Williams & Simpson 2014; based on **Milky Way Project** catalog (Simpson et al. 2013), which came from **Spitzer/GLIMPSE** (Churchwell et al. 2009, Benjamin et al. 2003), cf. Shenoy & Tan 2008 for discussion of HAC; **astroml.org** for machine learning advice/tools

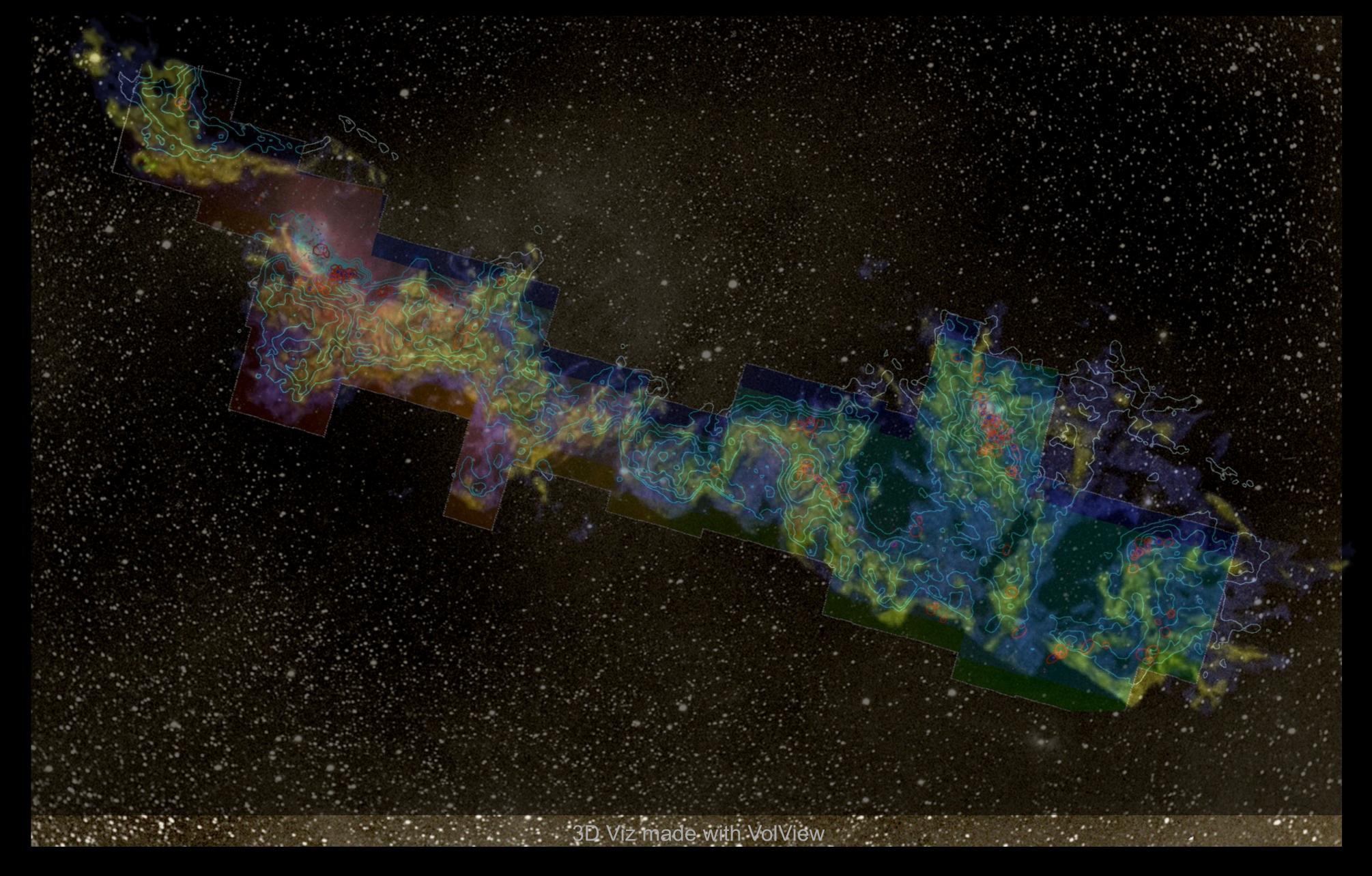


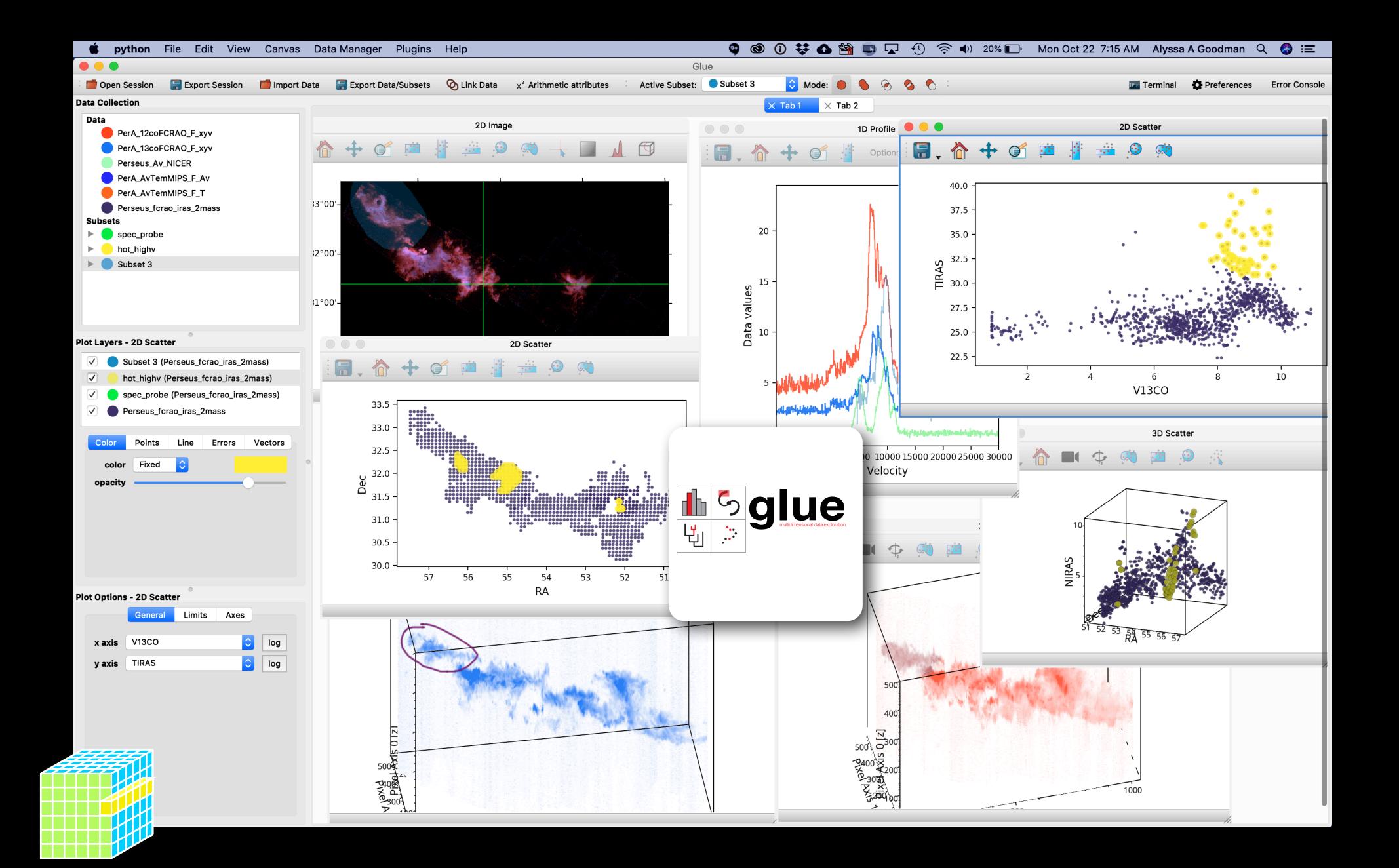
## BIG DATA, WIDE DATA

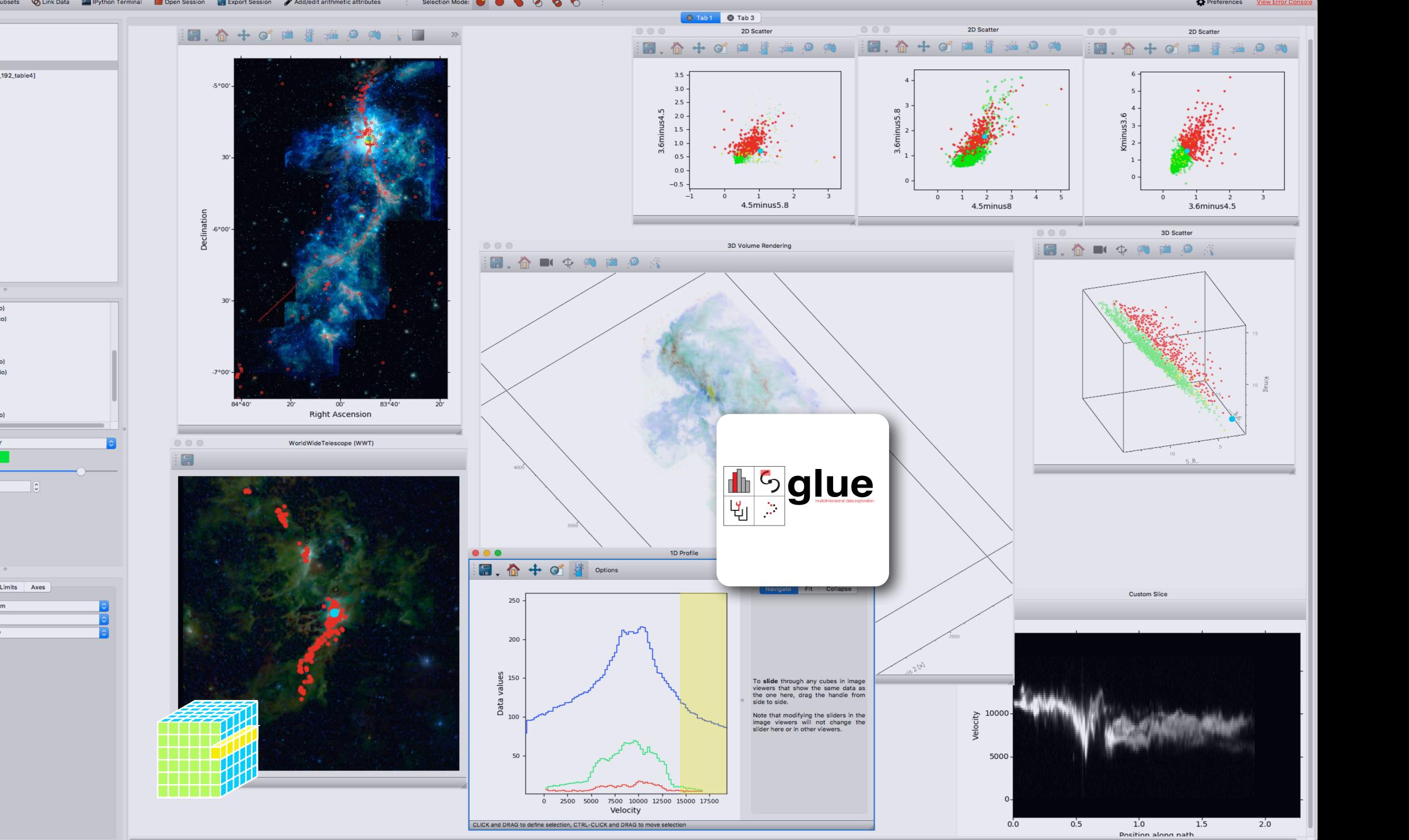


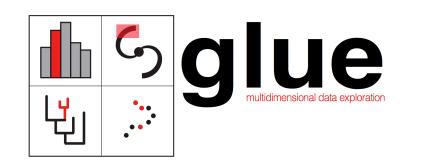
### C PLETE WIDE DATA mm peak (Enoch et al. 2006) sub-mm peak (Hatchell et al. 2005, Kirk et al. 2006) 13**CO** (Ridge et al. 2006) mid-IR IRAC composite from c2d data (Foster, Laakso, Ridge, et al.) Optical image (Barnard 1927)

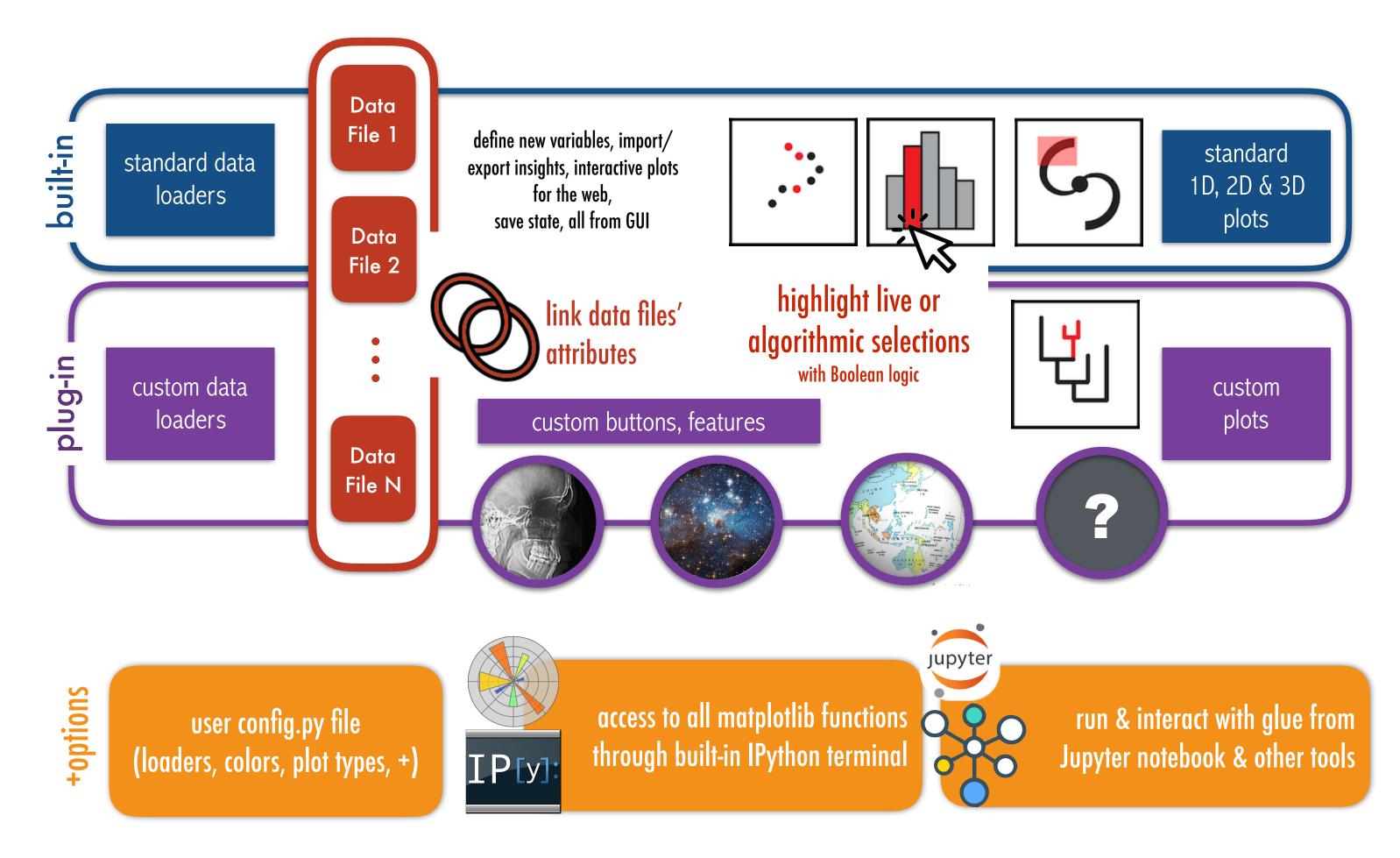






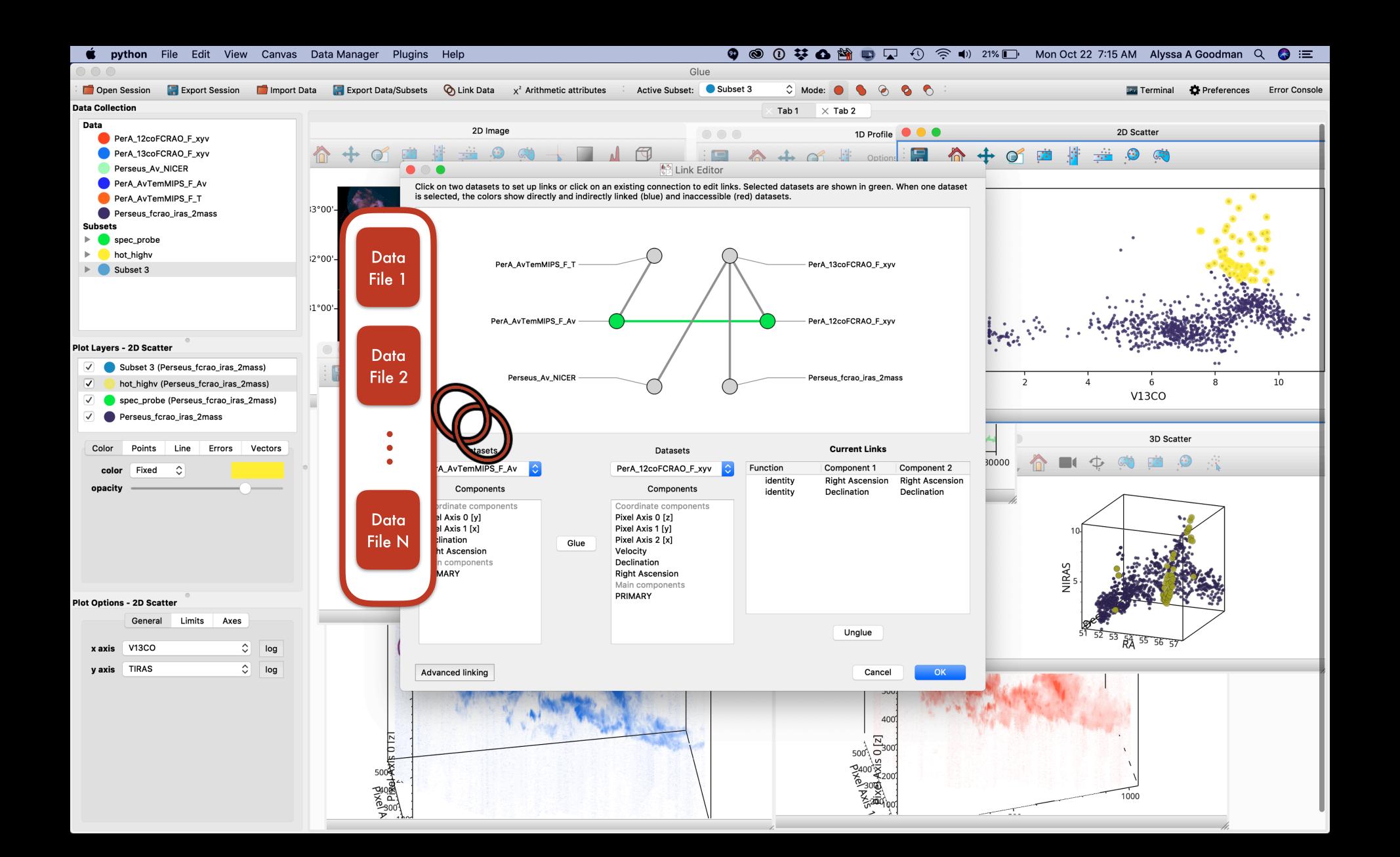




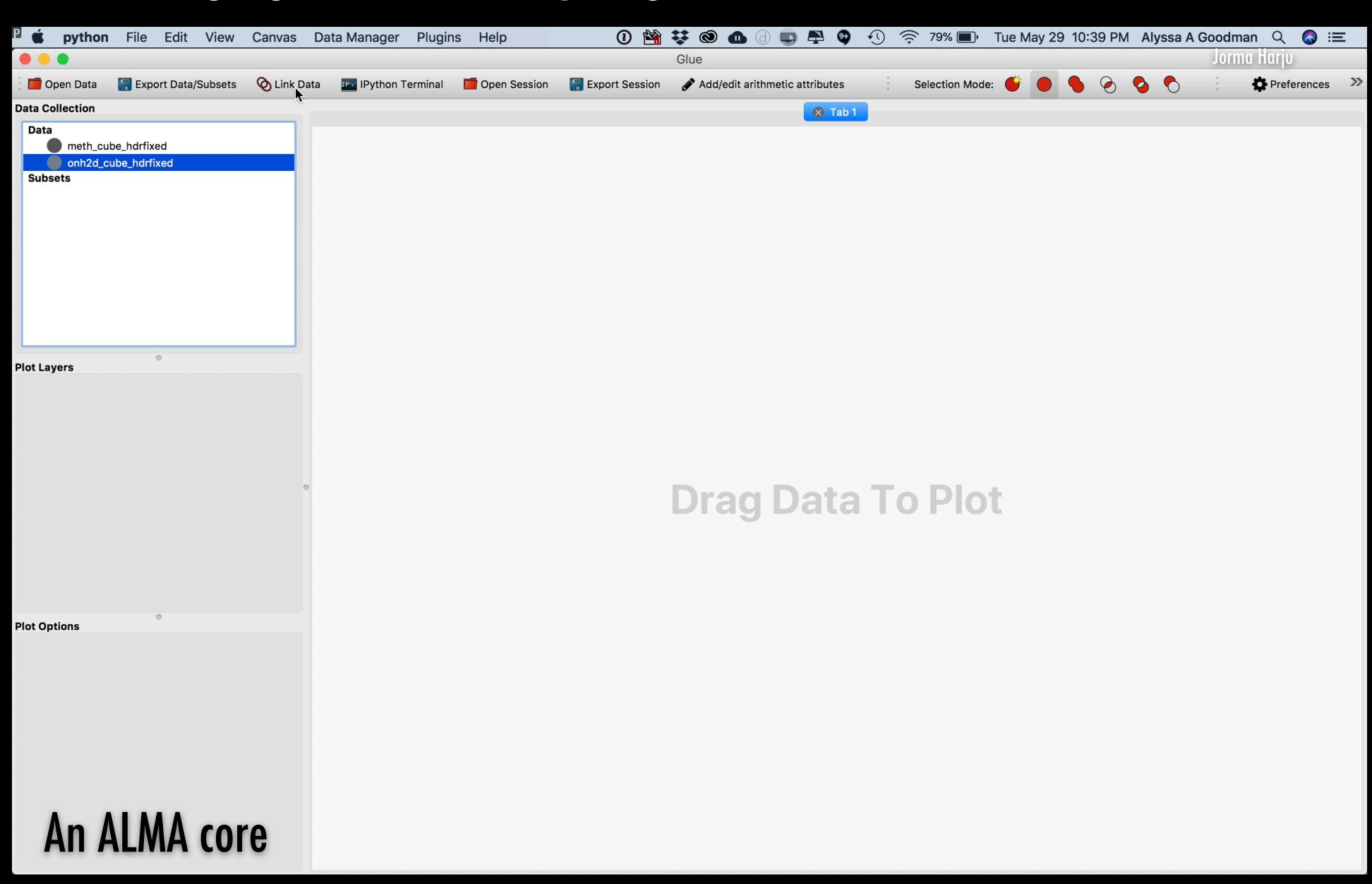


glueviz.org

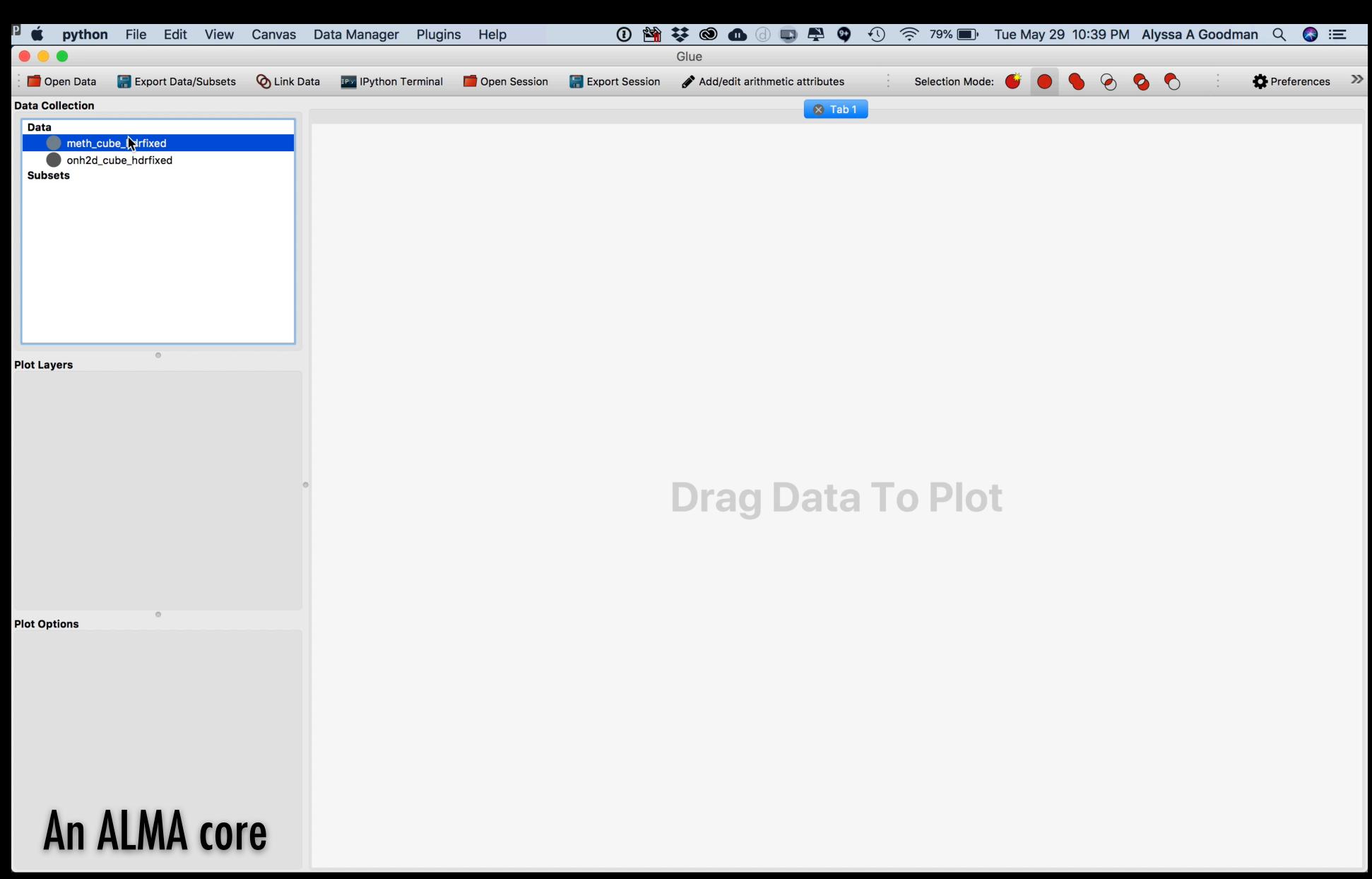
[demo]
[ALMA]
[publishing]



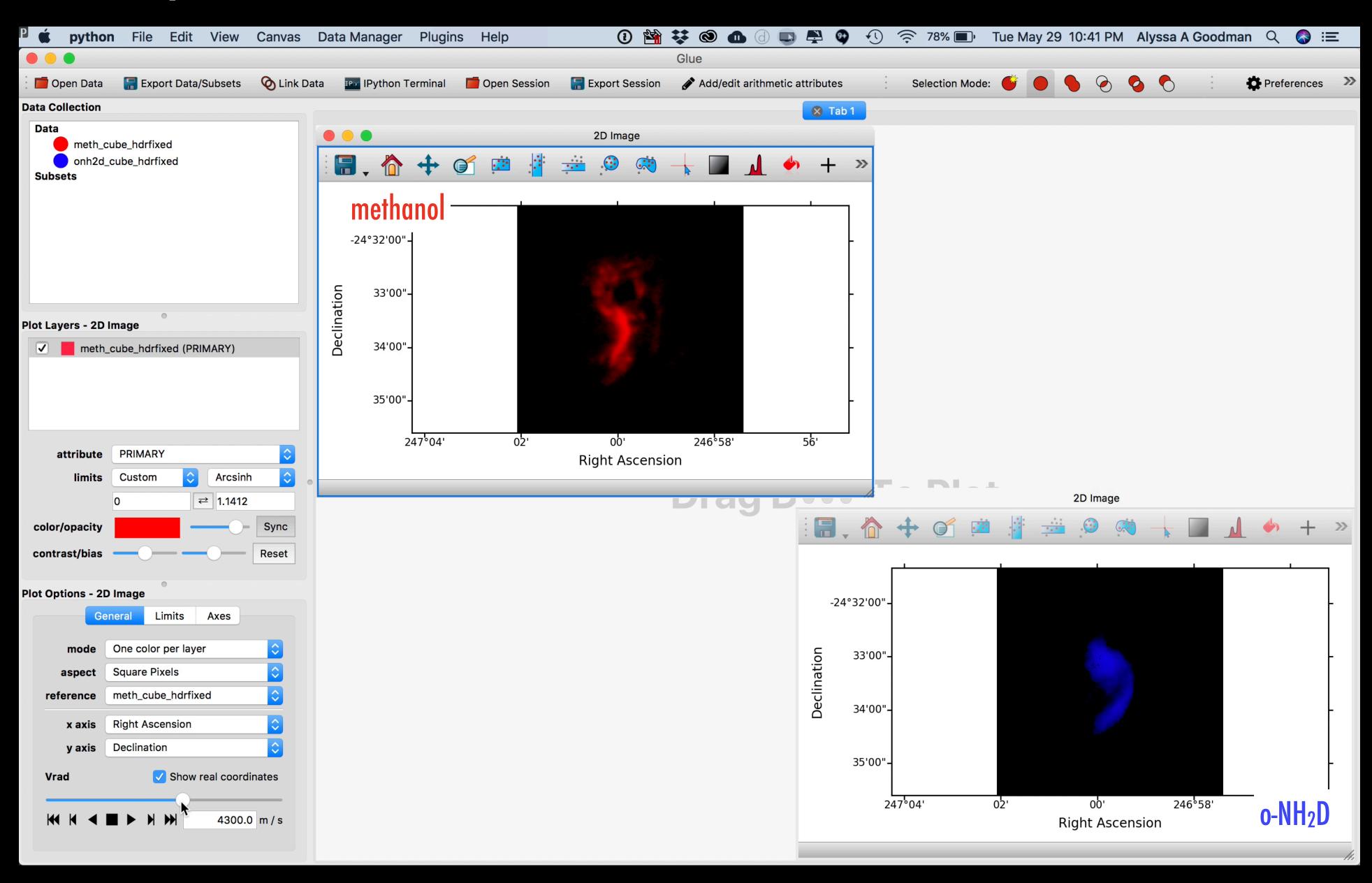
### "No merging of data sets—just glue them."



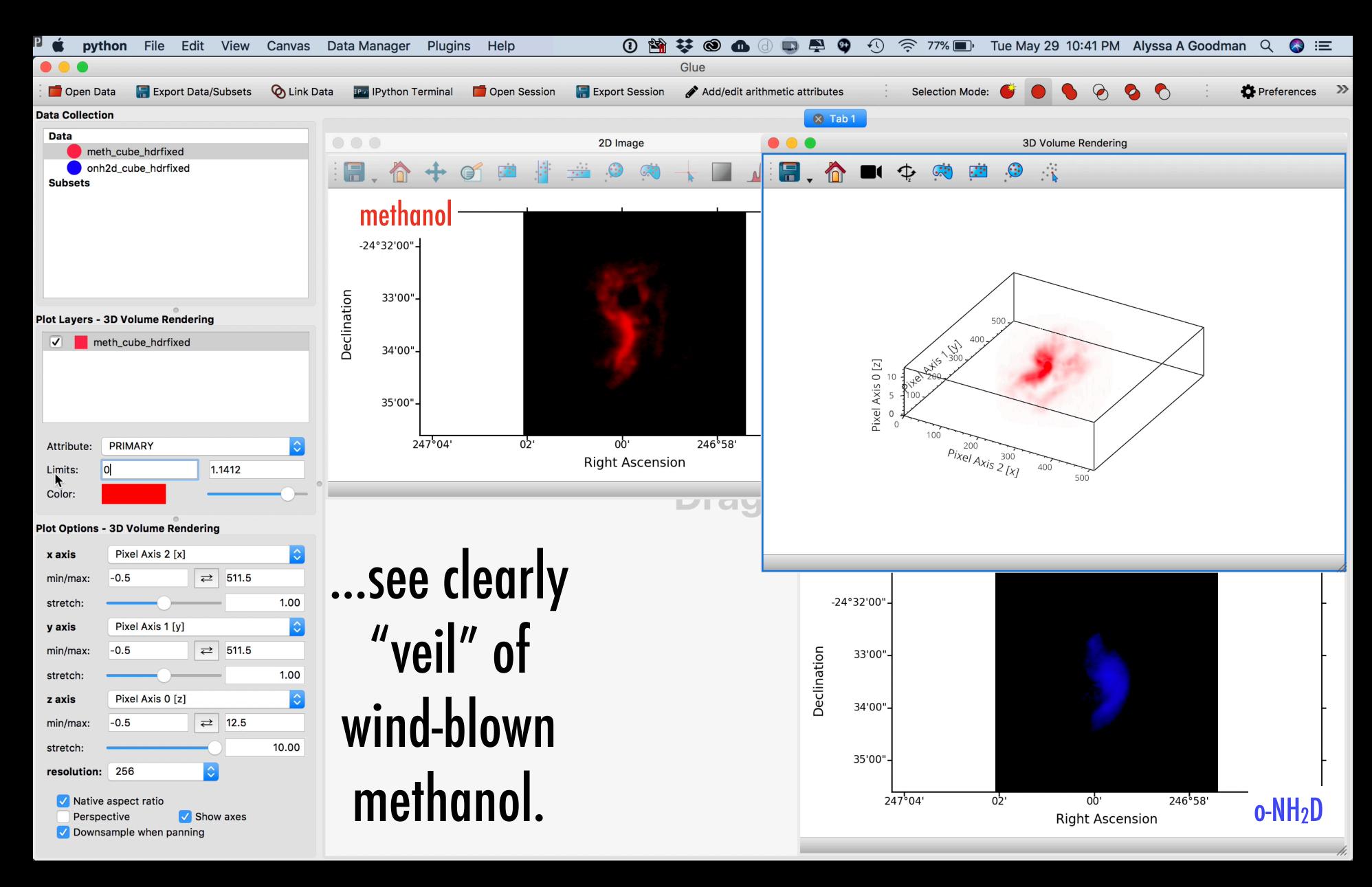
### Just drag to visualize, e.g. series of 2D "channel maps."



### Adjust so each tracer is a different color.

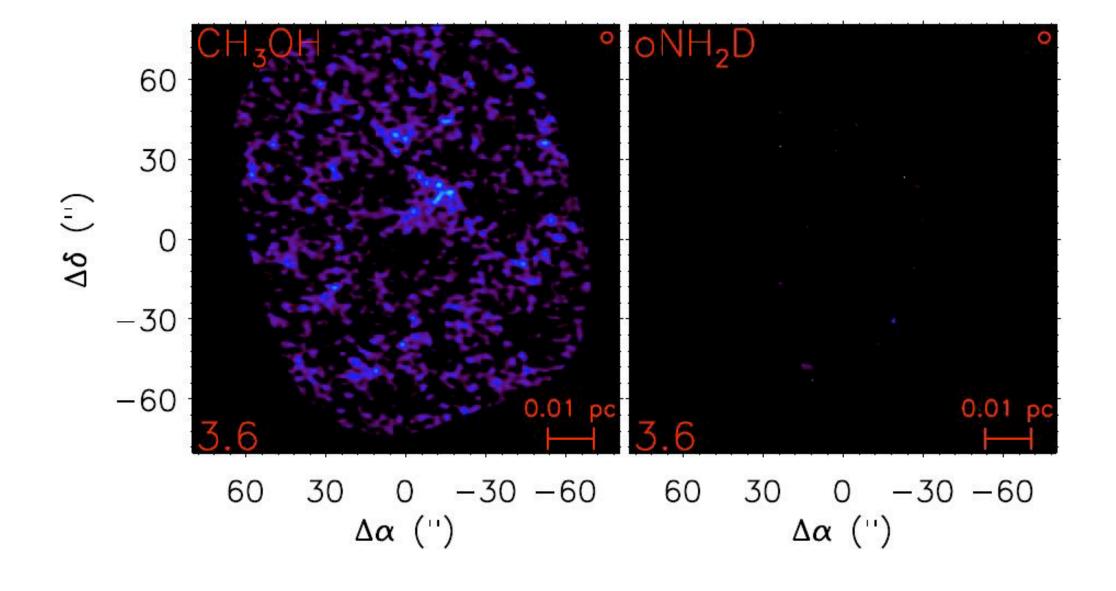


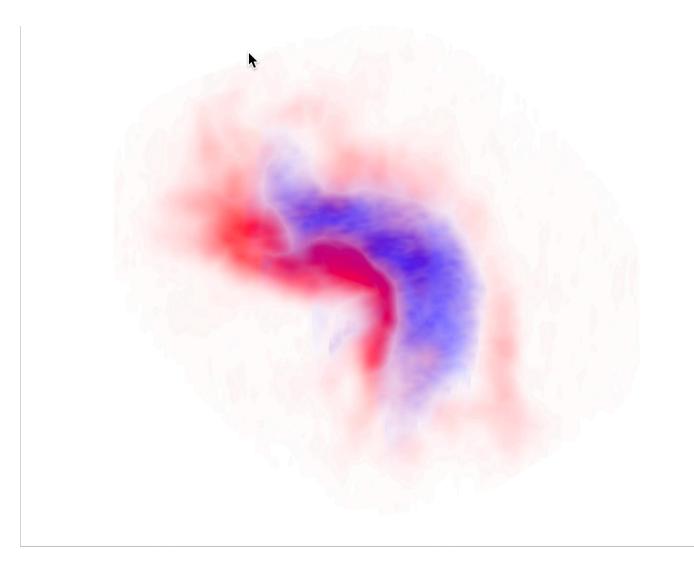
### Create 3D views...

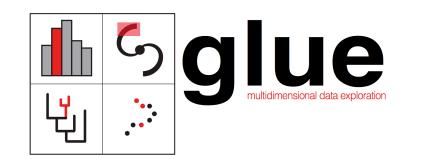


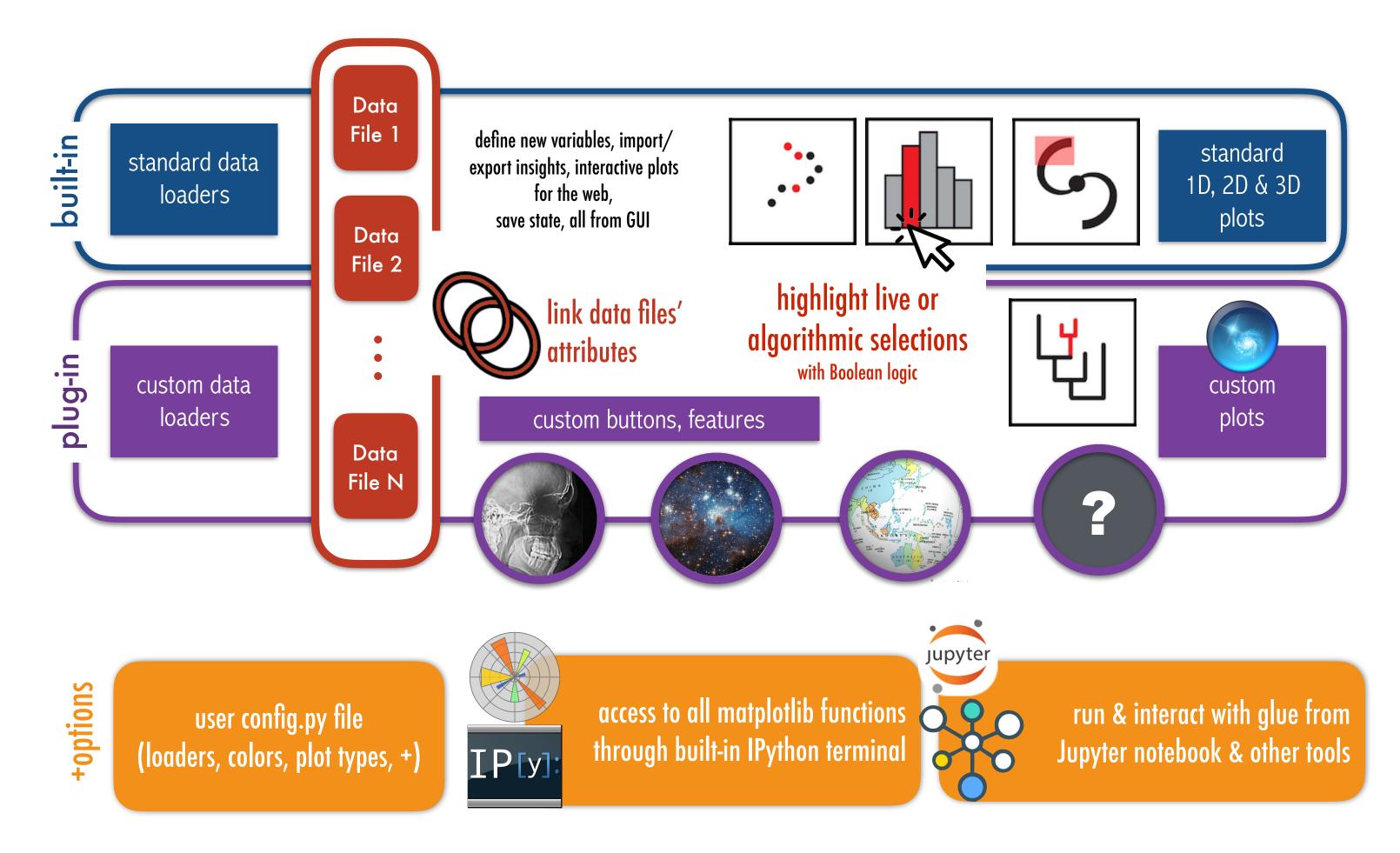
### Traditional Rainbow Channel maps

### glue





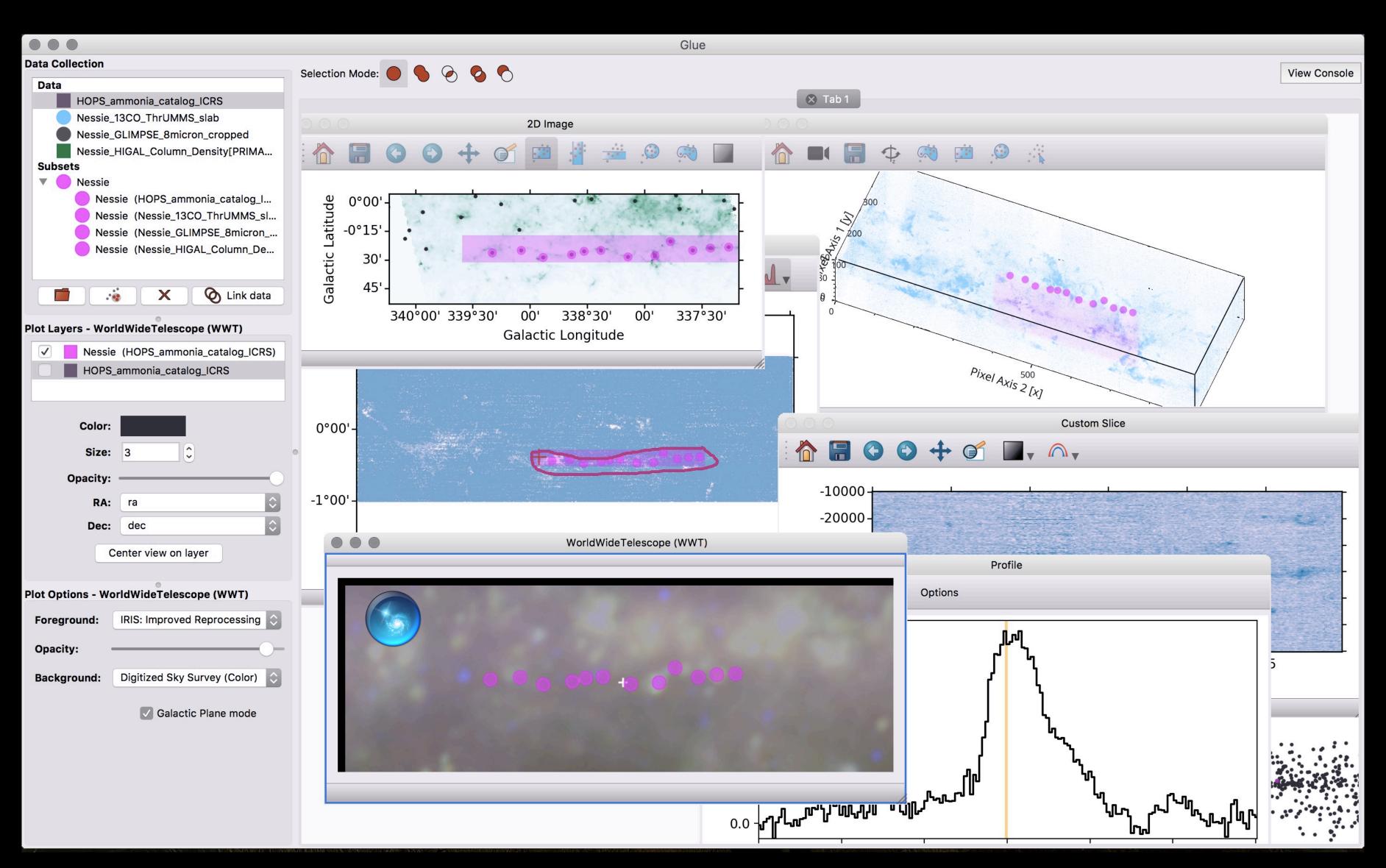




glueviz.org

[demo]
[ALMA]
[publishing]

### WorldWide Telescope as a plug-in to glue







# But... Publishing?

Au thorea FEATURED ARTICLES ABOUT PLANS BLOG FEED



(2)

21

#### TOBLIC MICOGIT DIVIL



o Settings

۶ Fork

Quicked

Word Count

🗪 41 C



#### The "Paper" of the Future



Alyssa Goodman, Josh Peek, Alberto Accomazzi, Chris Beaumont, Christine L. Borgman, How-Huan Hope Chen, Merce Crosas, Christopher Erdmann, August Muench, Alberto Pepe, Curtis Wong ♣ Add author ス Re-arrange authors

A 5-minute video demonsration of this paper is available at this YouTube link.

02

#### 1 Preamble

A variety of research on human cognition demonstrates that humans learn and communicate best when more than one processing system (e.g. visual, auditory, touch) is used. And, related research also shows that, no matter how technical the material, most humans also retain and process information best when they can put a narrative "story" to it. So, when considering the future of scholarly communication, we should be careful not to do blithely away with the linear narrative format that articles and books have followed for centuries: instead, we should enrich it.

Much more than text is used to commuicate in Science. Figures, which include images, diagrams, graphs, charts, and more, have enriched scholarly articles since the time of Galileo, and ever-growing volumes of data underpin most scientific papers. When scientists communicate face-to-face, as in talks or small discussions, these figures are often the focus of the conversation. In the best discussions, scientists have the ability to manipulate the figures, and to access underlying data, in real-time, so as to test out various what-if scenarios, and to explain findings more clearly. This short article explains—and shows with demonstrations—how scholarly "papers" can morph into long-lasting rich records of scientific discourse, enriched with deep data and code linkages, interactive figures, audio, video, and commenting.



Preamble Pof1 Collaborative authoring Comparison table Linking data Question 🝱 Dvn Zenodo Linking and executing .. Rho oph Better storytelling Audio Video Enhanced figures Interactivity Index 3d in 2d Nature screen shot Images in context

Deener easier citations





A Large Catalog of Accurate Distances to Local Molecular Clouds: The Gaia DR2 Edition

Catherine Zucker,<sup>1,\*</sup> Joshua S. Speagle,<sup>1,\*</sup> Edward F. Schlafly,<sup>2</sup> Gregory M. Green,<sup>3</sup> Douglas P. Finkbeiner,<sup>1</sup> Alyssa A. Goodman,<sup>1,4</sup> and João Alves<sup>4,5</sup>

<sup>1</sup>Harvard Astronomy, Harvard-Smithsonian Center for Astrophysics, 60 Garden St., Cambridge, MA 02138, USA

<sup>2</sup>Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, USA

<sup>3</sup>Kavli Institute for Particle Astrophysics and Cosmology, Physics and Astrophysics Building, 452 Lomita Mall, Stanford, CA 94305, USA

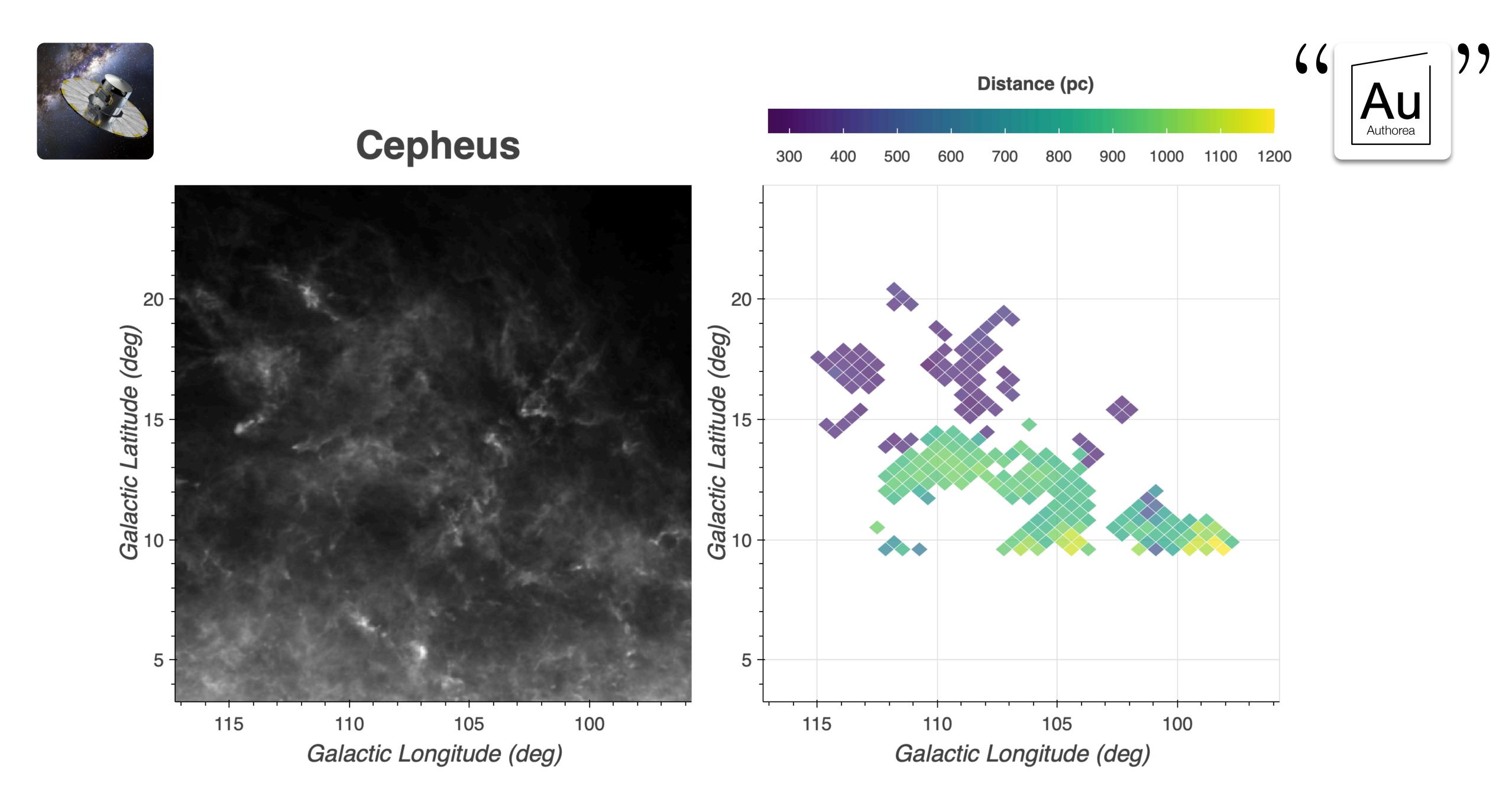
<sup>4</sup>Radcliffe Institute for Advanced Study, Harvard University, 10 Garden St, Cambridge, MA 02138

<sup>5</sup>University of Vienna, Department of Astrophysics, Türkenschanzstraße 17, 1180 Vienna, Austria

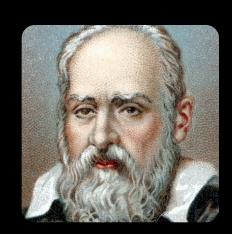
#### ABSTRACT

We present a uniform catalog of accurate distances to local molecular clouds informed by the Gaia DR2 data release. Our methodology builds on that of Schlafly et al. (2014). First, we infer the distance and extinction to stars along sightlines towards the clouds using optical and near-infrared photometry. When available, we incorporate knowledge of the stellar distances obtained from Gaia DR2 parallax measurements. We model these per-star distance-extinction estimates as being caused by a dust screen with a 2-D morphology derived from Planck at an unknown distance, which we then fit for using a nested sampling algorithm. We provide updated distances to the Schlafly et al. (2014) sightlines towards the Dame et al. (2001) and Magnani et al. (1985) clouds, finding good agreement with the earlier work. For a subset of 27 clouds, we construct interactive pixelated distance maps to further study detailed cloud structure, and find several clouds which display clear distance gradients and/or are comprised of multiple components. We use these maps to determine robust average distances to these clouds. The characteristic combined uncertainty on our distances is  $\approx 5-6\%$ , though this can be higher for clouds at farther distances, due to the limitations of our single-cloud model.

Keywords: ISM: clouds, ISM: dust, extinction, stars: distances, methods: statistical



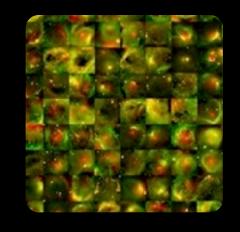
Interactive Gaia-reliant figure from <a href="Zucker et al. 2019">Zucker et al. 2019</a>— <a href="https://faun.rc.fas.harvard.edu/czucker/Paper\_Figures/Cepheus.html">https://faun.rc.fas.harvard.edu/czucker/Paper\_Figures/Cepheus.html</a>

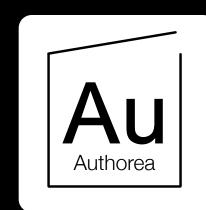


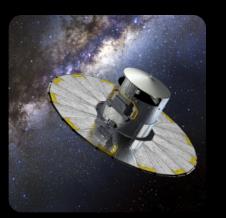
















 $\equiv$  MENU



#### TEN QUESTIONS TO ASK WHEN CREATING A VISUALIZATION

#### The 10 Questions

- 1. **Who** | Who is your audience? How expert will they be about the subject and/or display conventions?
- 2. **Explore-Explain** | Is your goal to explore, document, or explain your data or ideas, or a combination of these?
- 3. Categories | Do you want to show or explore pre-existing, known, human-interpretable, categories?
- 4. Patterns | Do you want to identify new, previously unknown or undefined patterns?
- 5. **Predictions & Uncertainty** | Are you making a comparison between data and/or predictions? Is representing uncertainty a concern?
- 6. **Dimensions** | What is the intrinsic number of dimensions (not necessarily spatial) in your data, and how many do you want to show at once?
- 7. **Abstraction & Accuracy** | Do you need to show all the data, or is summary or abstraction OK?
- 8. **Context & Scale** | Can you, and do you want to, put the data into a standard frame of reference, coordinate system, or show scale(s)?
- 9. **Metadata** | Do you need to display or link to non-quantitative metadata? (including captions, labels, etc.)
- 10. **Display Modes** | What display modes might be used in experiencing your display?



Join the 10QViz Conversation!



To learn more about this site, please visit the **About** page.

To read an in-process manuscript giving the scholarship behind the recommendations on this site, see Coltekin & Goodman 2018.

# Bonus anyone?



Infoviz

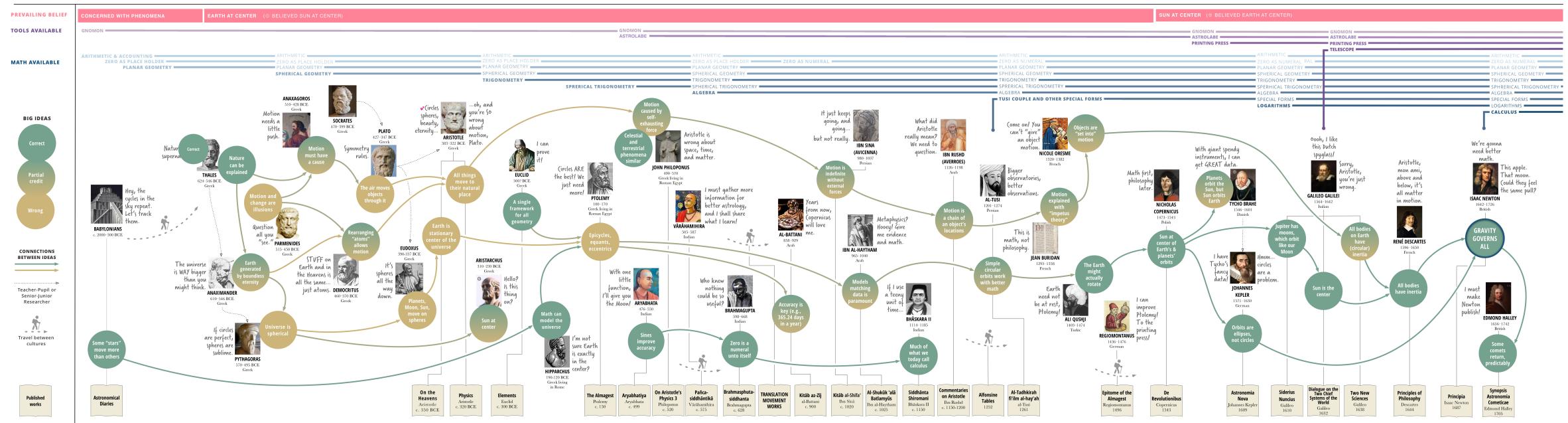


3D selection



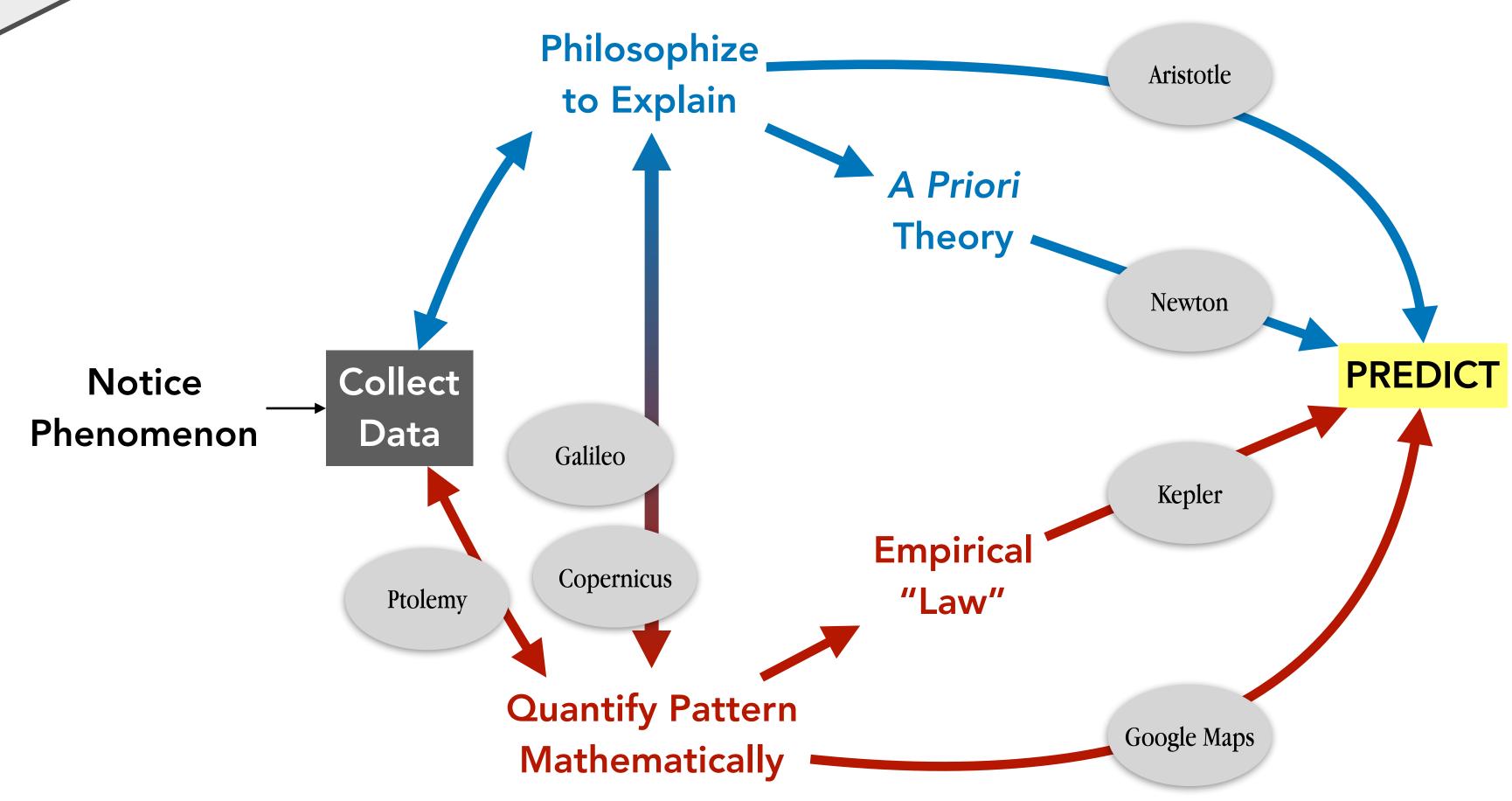
#### The Path to Newton



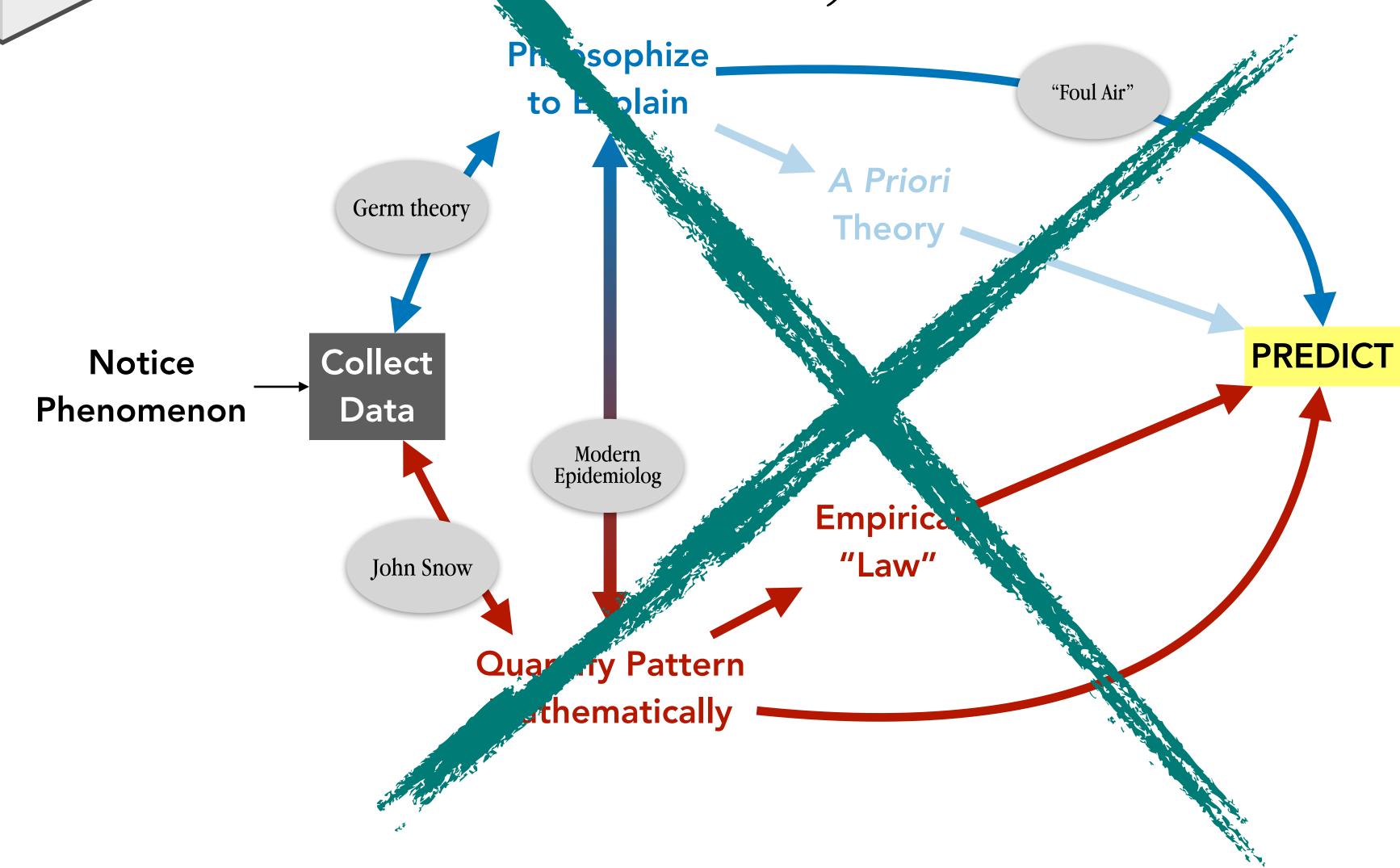




# Version 1 (Terrible)







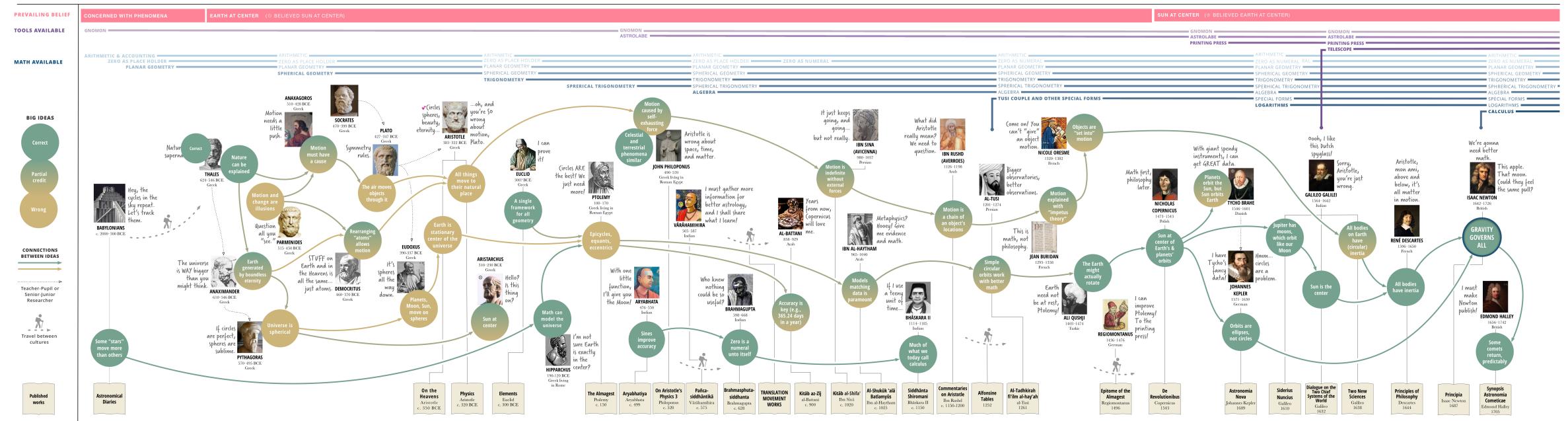
Version 2 . . . Simplify!



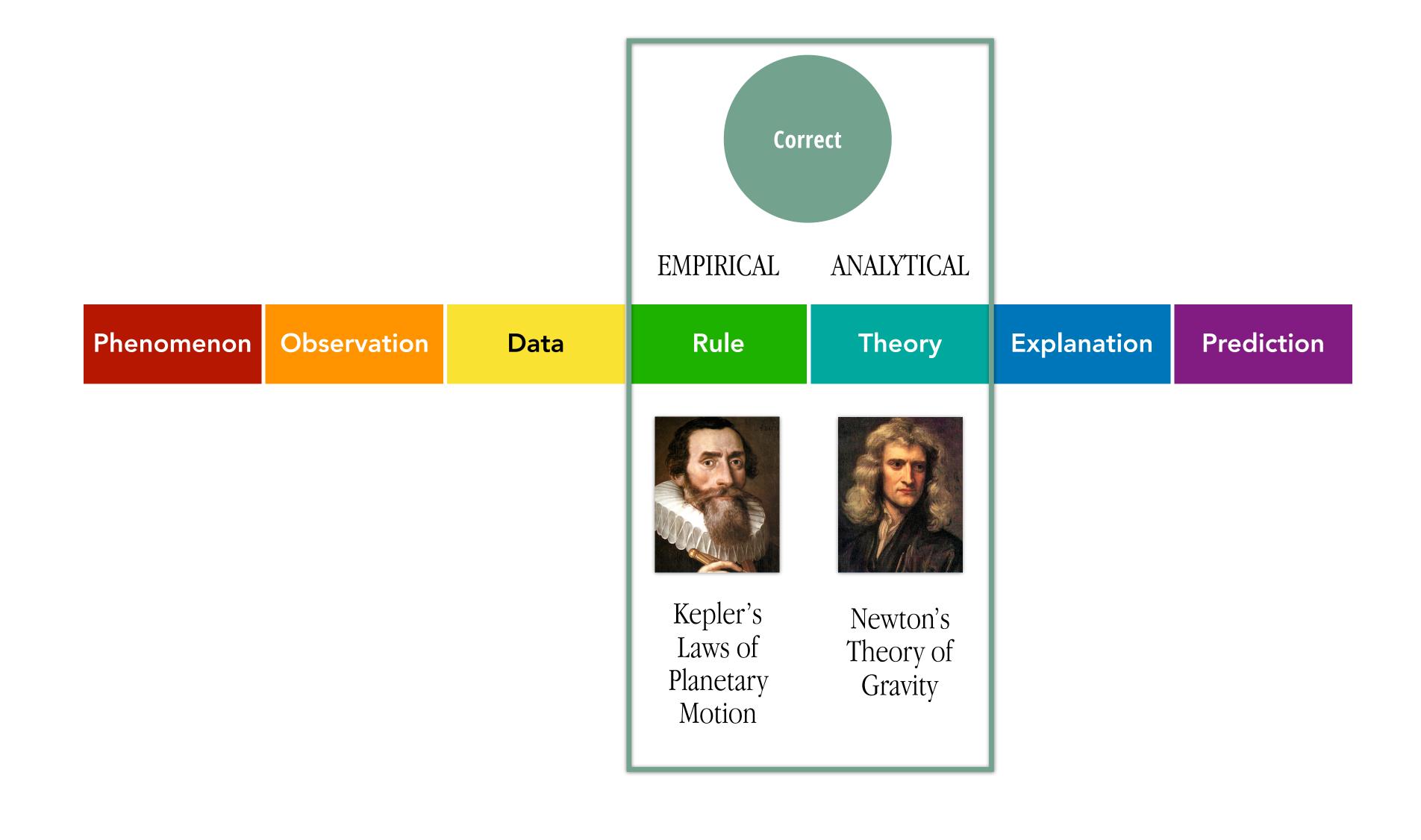
Phenomenon Observation\* Data Rule Theory Explanation Prediction

#### The Path to Newton



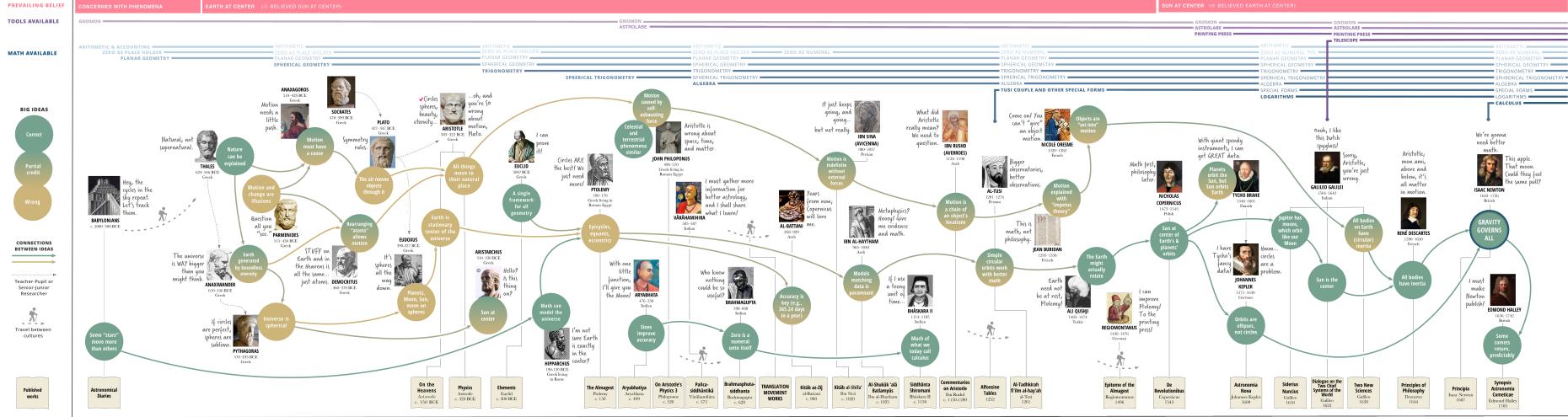


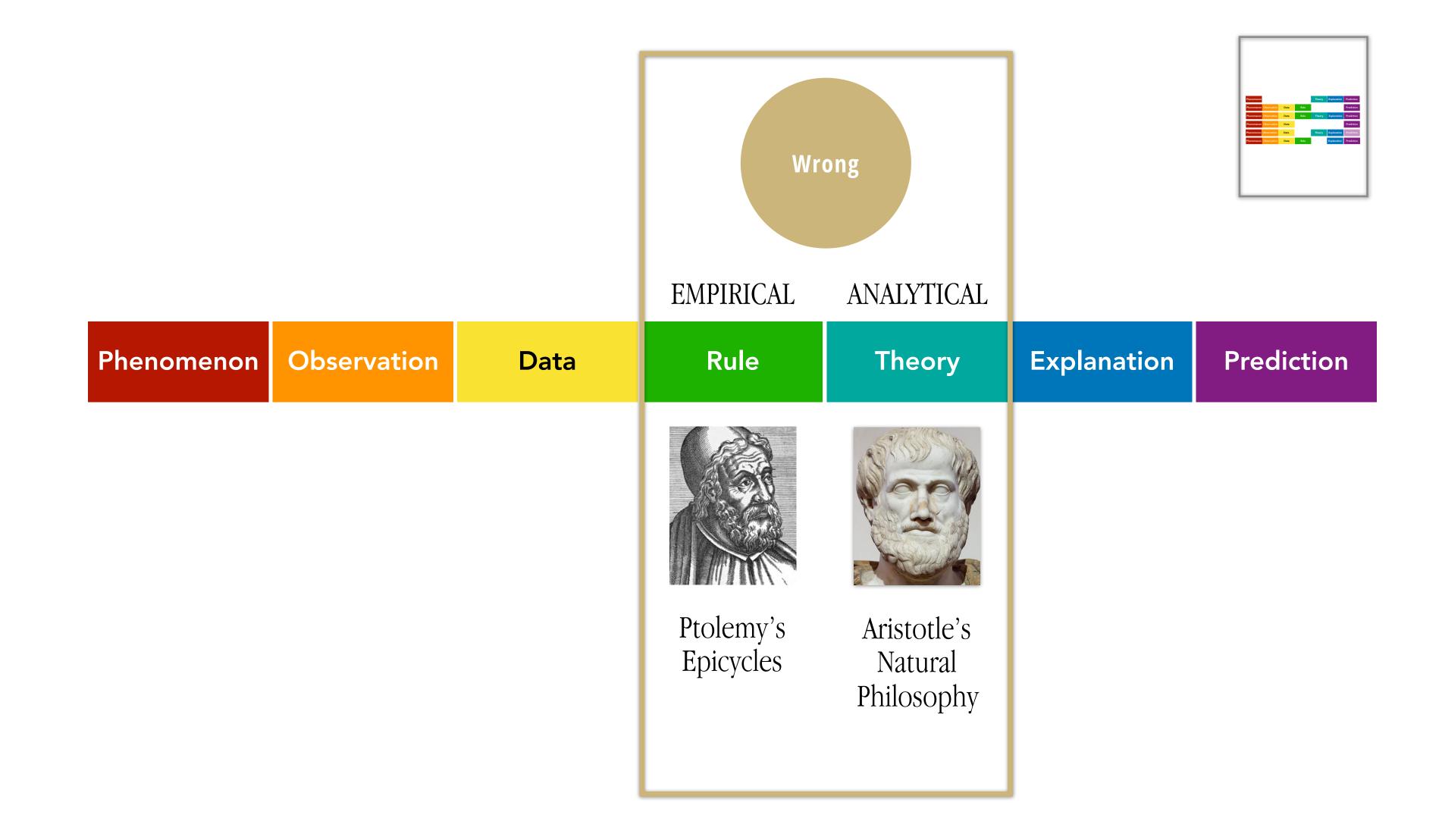
© Harvard University, created by Alyssa Goodman, Jais Brohinsky, Drew Lichtenstein & Katie Peek, re-use is allowed, with attribution, version 1, 201

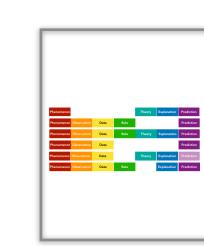


Phenomenon Observation Data Rule Theory Explanation Prediction

### The Path to Newton PREVAILING BELIEF CONCERNED WITH PHENOMENA CONCERNE



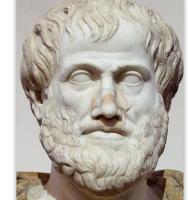






#### Babylonians

Phenomenon	Observation	Data	Rule
------------	-------------	------	------

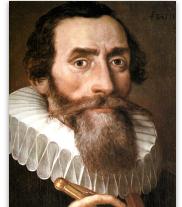


Aristotle

Phenomenon Observation

Theory Explanation

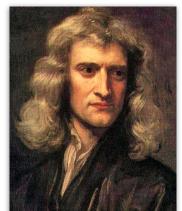
Prediction



Kepler

Phenomenon Observation Data Rule

Prediction



Newton

Phenomenon Observation Data Rule Theory Explanation Prediction



EMPIRICAL STATISTICAL

 Phenomenon
 Observation\*
 Data
 Rule
 Theory
 Explanation
 Prediction



Mendel's Rules



Darwin's
Theory of
Evolution

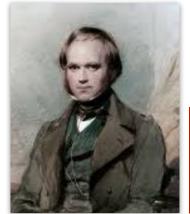




Mendel

Phenomenon Observation Data Rule

Prediction



Darwin

Phenomenon Observation Data

Theory

Explanation

Prediction



Mendel

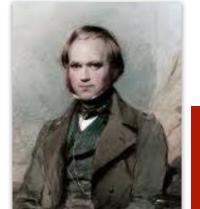
Phenomenon

Observation

Data

Rule

Prediction



Darwin

Phenomenon

**Observation** 

Data

Theory

**Explanation** 

Prediction

## "YOUR" WORLD (BIOLOGY)

## MY WORLD (PHYSICS)



Kepler

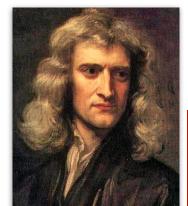
Phenomenon

**Observation** 

Data

Rule

Prediction



Newton

Phenomenon

Observation

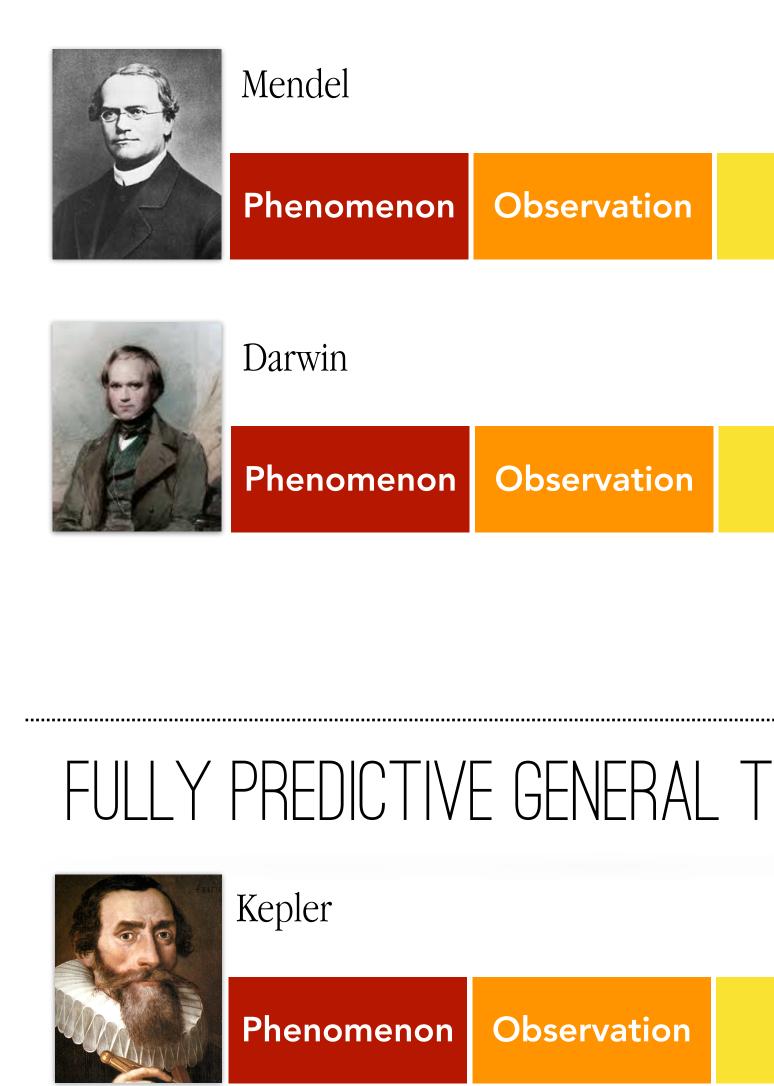
Data

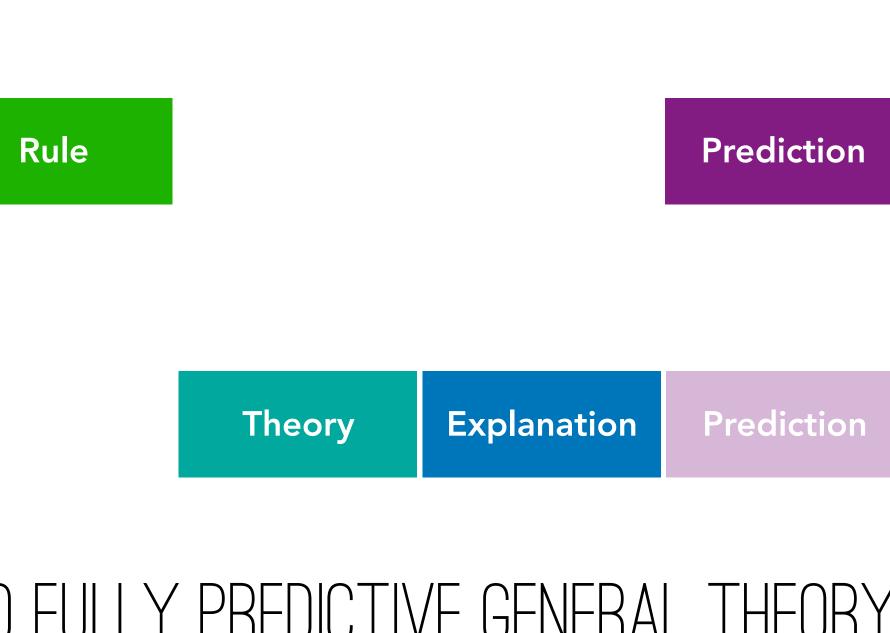
Rule

Theory

**Explanation** 

Prediction





## NO FULLY PREDICTIVE GENERAL THEORY

### FULLY PREDICTIVE GENERAL THOERY

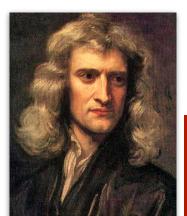


Rule Data

Data

Data

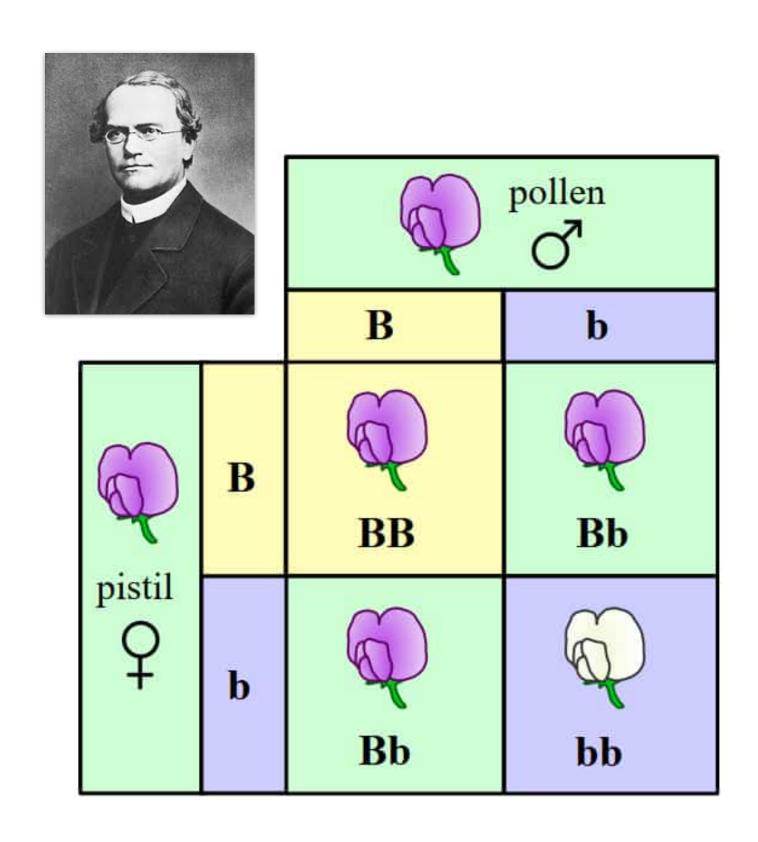
Prediction

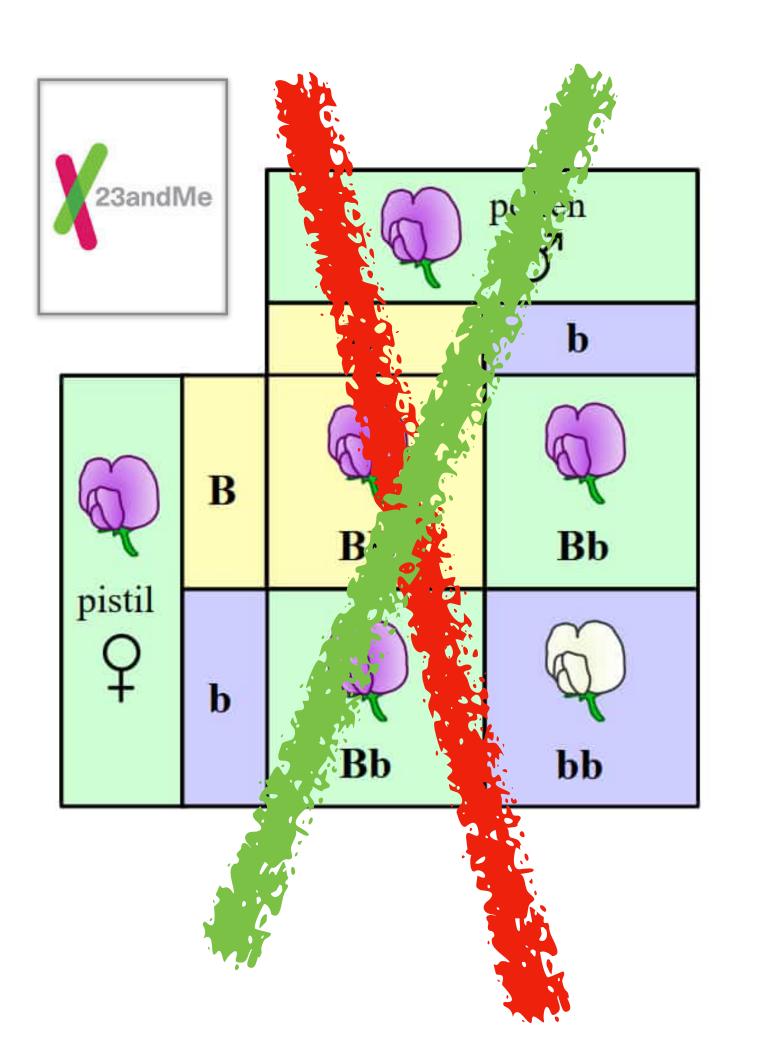


Newton

Phenomenon Observation **Explanation** Rule Theory Prediction Data

### NO FULLY PREDICTIVE GENERAL THEORY





# THE FUTURE OF THE FUTURE

### 20th century



#### 21st century?





## Coming next: glue in the browser

