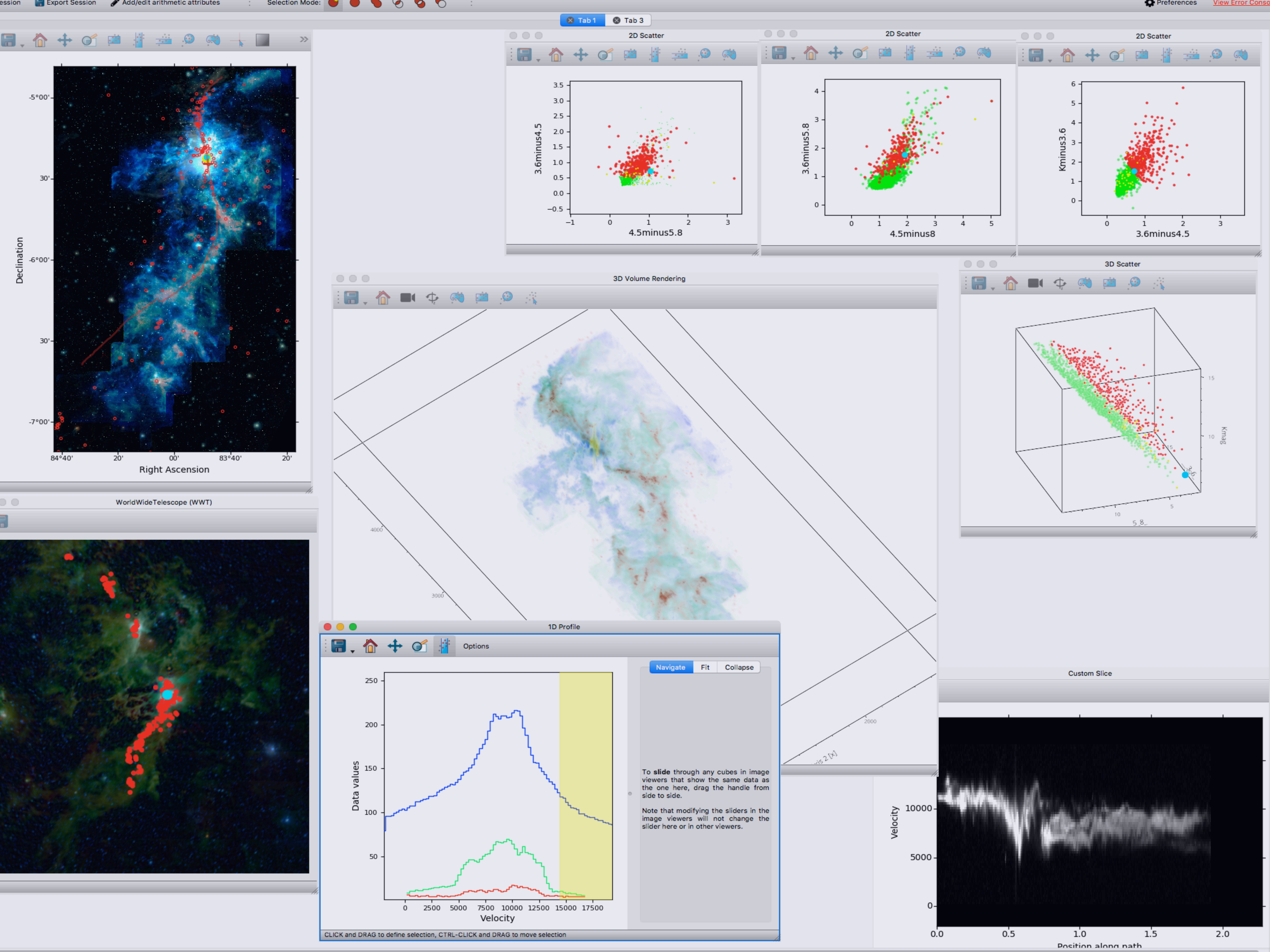
Visualization & the Universe

How and why astronomers, doctors, and you need to work together to understand the world around us...

Alyssa A. Goodman
Harvard-Smithsonian Center for Astrophysics + Radcliffe Insitute for Advanced Study
@AlyssaAGoodman


Export Session
Add/edit arithmetic attributes
Selection Mode:
Preferences
View Error Console
Tab 1
Tab 3
Declination
Right Ascension
2D Scatter
3.6minus4.5
4.5minus5.8
4.5minus8
3.6minus5.8
Kminus3.6
3D Volume Rendering
3D Scatter
WorldWideTelescope (WWT)
1D Profile
Options
Navigate
Fit
Collapse
Data values
Velocity
To slide through any cubes in image viewers that show the same data as the one here, drag the handle from side to side.
Note that modifying the sliders in the image viewers will not change the slider here or in other viewers.
CLICK and DRAG to define selection, CTRL-CLICK and DRAG to move selection
Custom Slice
Velocity
Position along path

What anyone can do right now

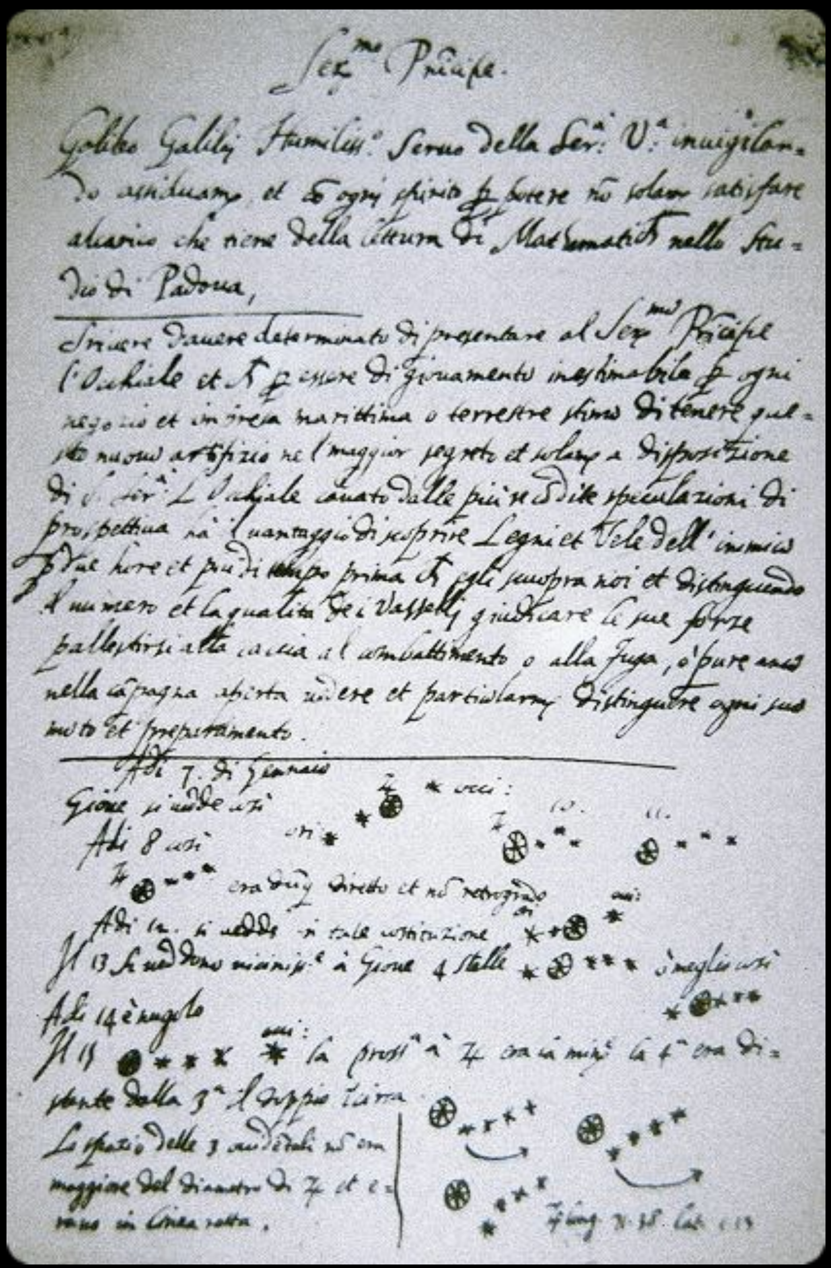
What computer scientists do

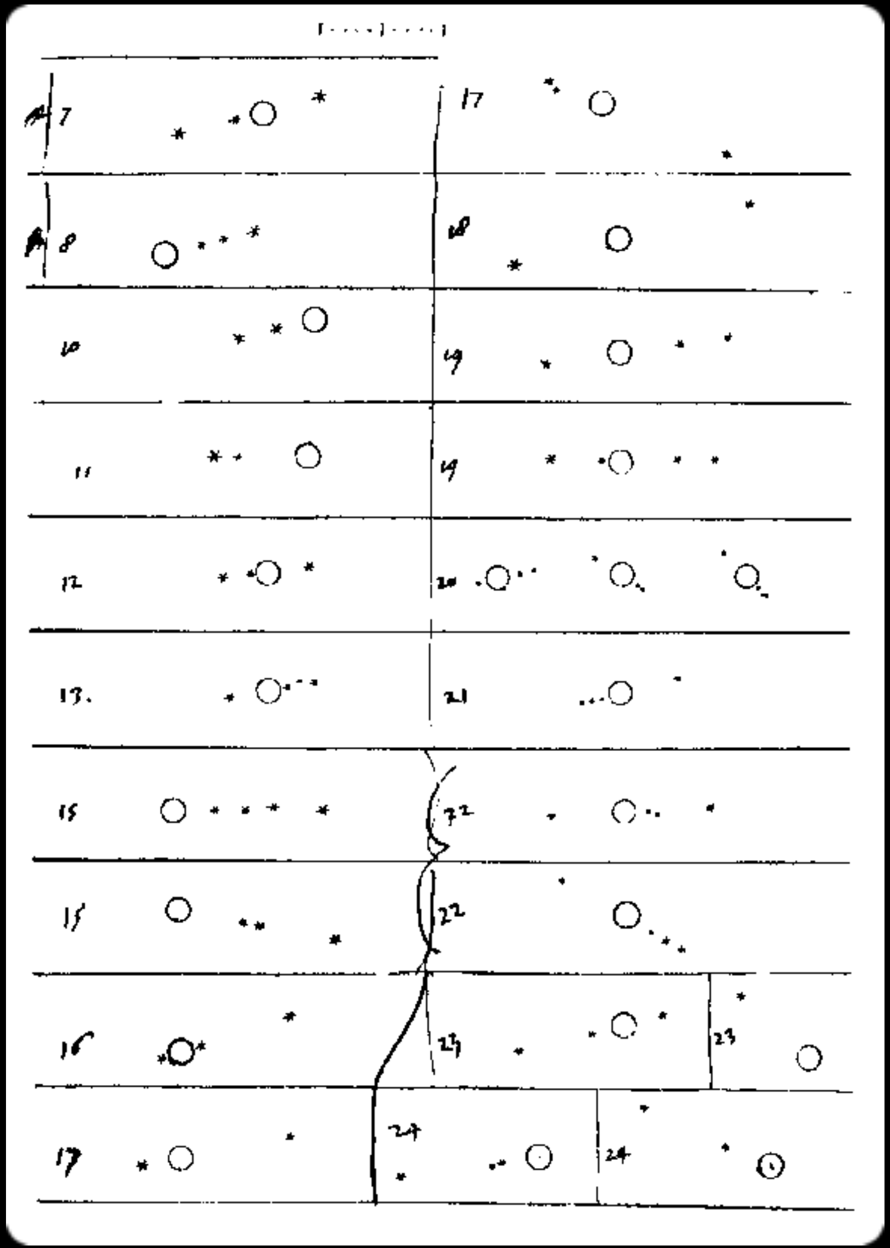
What scientists do

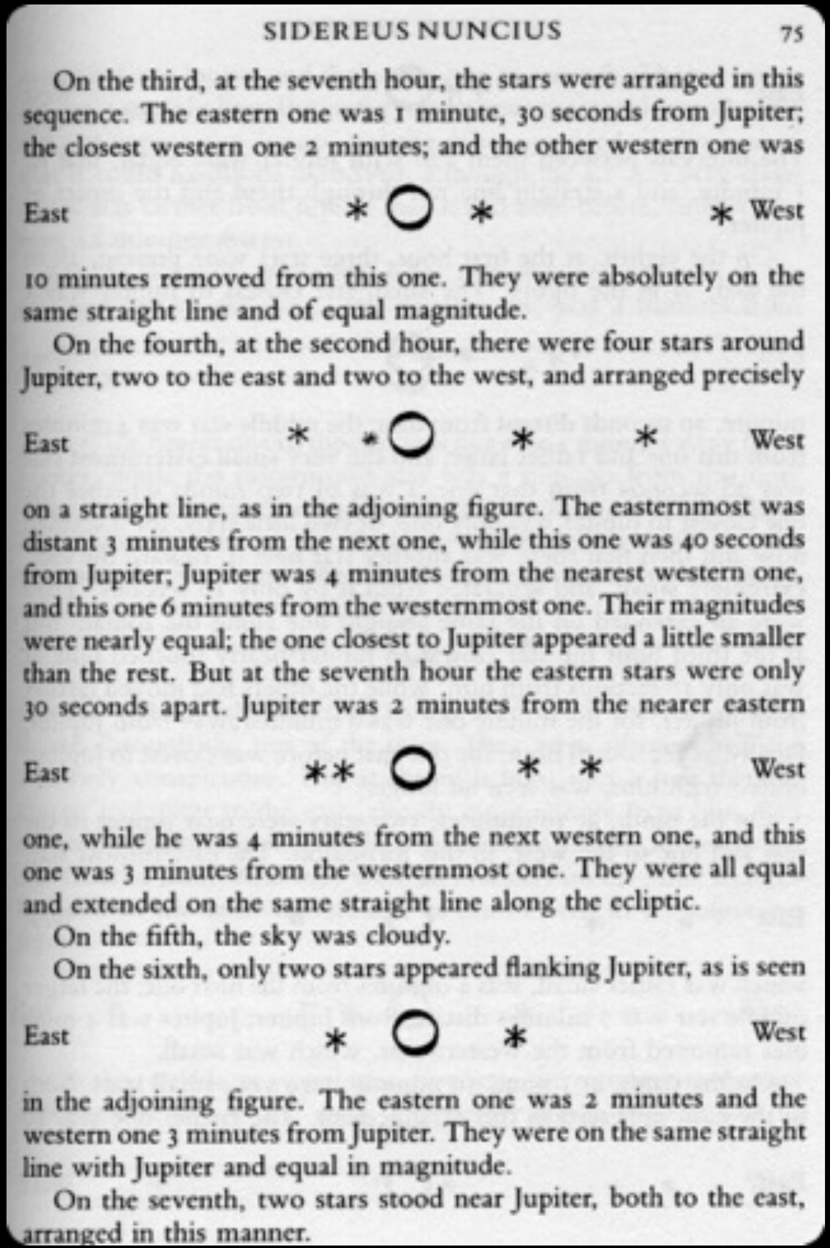
What we should do?

3D PDF	PotF	10QViz	glupyter	
2009	2015	2018	2019	+ more, together

1610: Galileo Discovers Jupiter's Moons







On the third, at the seventh hour, the stars were arranged in this sequence. The eastern one was 1 minute, 30 seconds from Jupiter; the closest western one 2 minutes; and the other western one was

East * O * * West

10 minutes removed from this one. They were absolutely on the same straight line and of equal magnitude.

On the fourth, at the second hour, there were four stars around Jupiter, two to the east and two to the west, and arranged precisely

East * * O * * West

on a straight line, as in the adjoining figure. The easternmost was distant 3 minutes from the next one, while this one was 40 seconds from Jupiter; Jupiter was 4 minutes from the nearest western one, and this one 6 minutes from the westernmost one. Their magnitudes were nearly equal; the one closest to Jupiter appeared a little smaller than the rest. But at the seventh hour the eastern stars were only 30 seconds apart. Jupiter was 2 minutes from the nearer eastern

East ** O * * West

one, while he was 4 minutes from the next western one, and this one was 3 minutes from the westernmost one. They were all equal and extended on the same straight line along the ecliptic.

On the fifth, the sky was cloudy.

On the sixth, only two stars appeared flanking Jupiter, as is seen

East * O * West

in the adjoining figure. The eastern one was 2 minutes and the western one 3 minutes from Jupiter. They were on the same straight line with Jupiter and equal in magnitude.

On the seventh, two stars stood near Jupiter, both to the east, arranged in this manner.

Notes for & re-productions of Siderius Nuncius

WorldWide Telescope: Explaining Galileo's Discovery

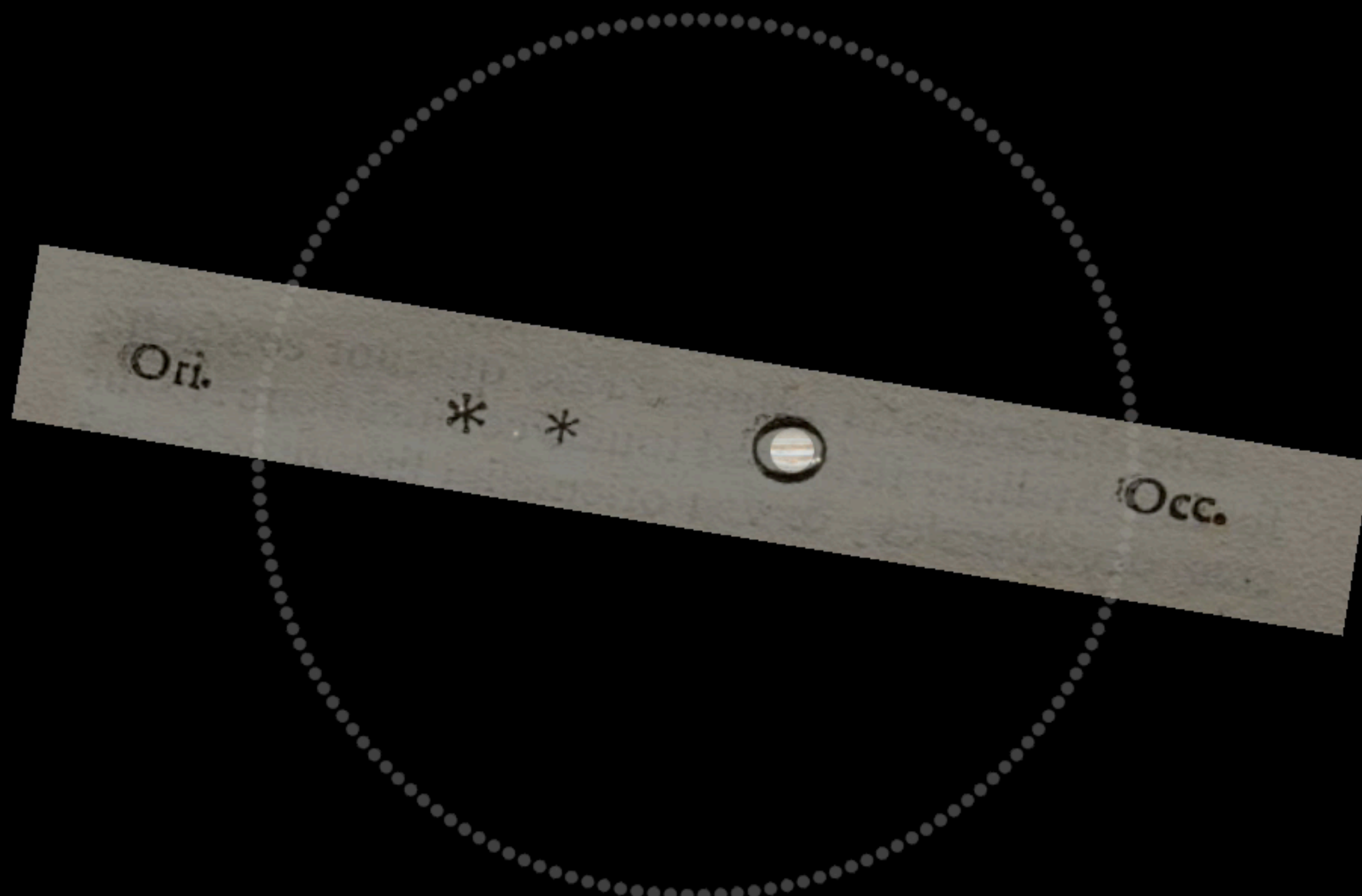
GALILEO'S "NEW ORDER"

Created by Alyssa Goodman, Curtis Wong and Pat Udomprasert,
with advice from Owen Gingerich and David Malin

Galileo's New Order, A WorldWide Telescope Tour by Goodman, Wong & Udomprasert 2010
WWT Software Wong (inventor, MS Research), Fay (architect, MS Research), et al., now open source, hosted by AAS
see wwtambassadors.org for more on WWT Outreach

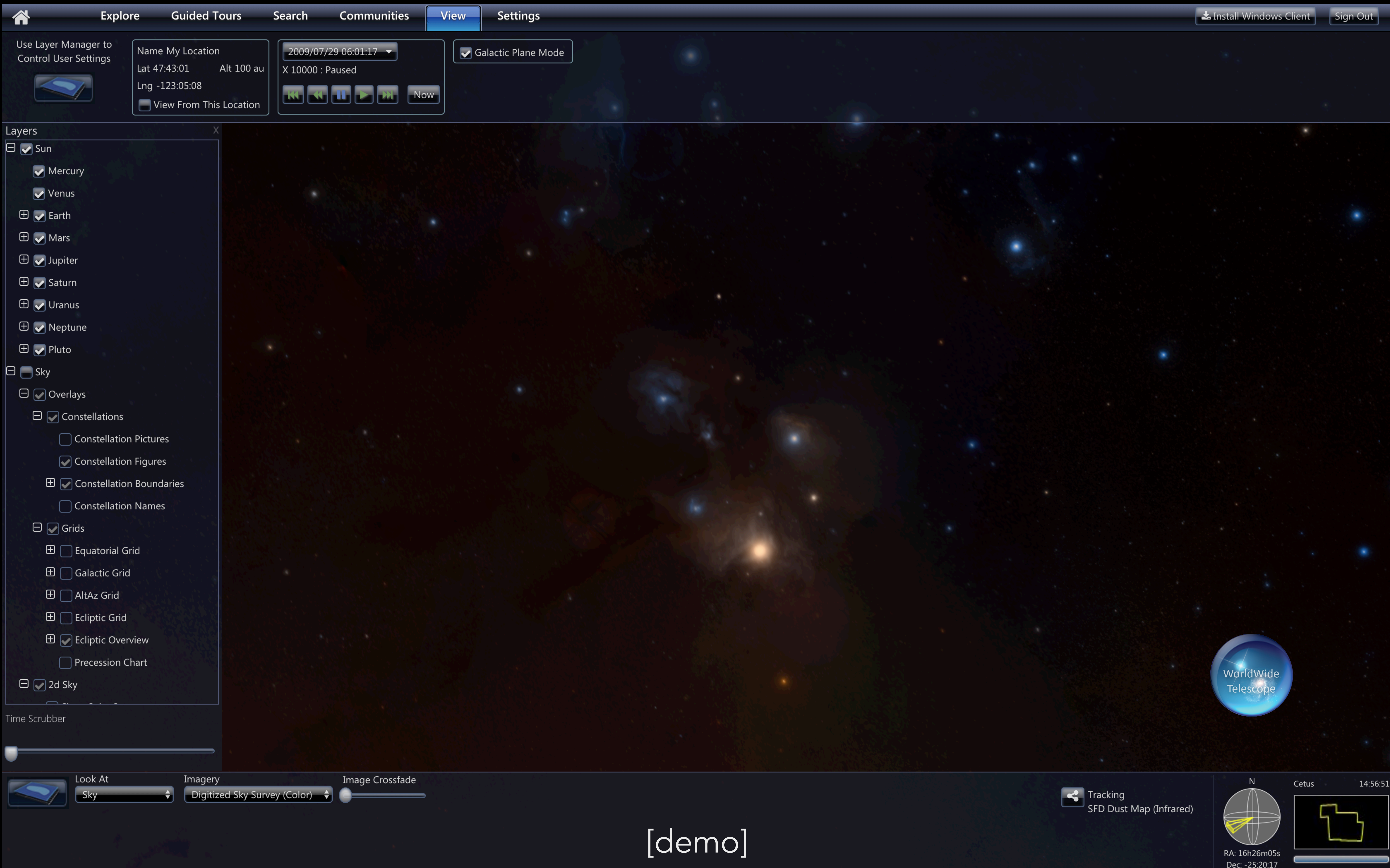
WorldWide Telescope: Explaining Galileo's Discovery

January 11, 1610



Galileo's New Order, A WorldWide Telescope Tour by Goodman, Wong & Udomprasert 2010
WWT Software Wong (inventor, MS Research), Fay (architect, MS Research), et al., now open source, hosted by AAS
see wwtambassadors.org for more on WWT Outreach

WorldWide Telescope: Exploring the Sky at Many Wavelengths



[demo]

Visualization in Computer Science

VIS 2018 SCHEDULE

WEDNESDAY, 24 OCTOBER

Tracks: VAST (Conv 1, Sec C) · INFOVIS (Conv 1, Sec D) · SCIVIS (Estrel A+B) · VIS (Estrel C) · VIS (Room IV)

9:00 AM – 10:40 AM
- VAST: Ensemble and Provenance — DS
- INFOVIS: Time
- SCIVIS: Biomedical Visualization — BA
- VIS (Estrel C): VAST: Text — DS BA
- VIS (Room IV): VisAP: Arts Program Session 1: Arts and Society

10:40 AM BREAK

11:00 AM – 12:40 PM
- VAST: Applications — AP
- INFOVIS: Graphs & Trees
- SCIVIS: Volume Visualization
- VIS (Estrel C): InfoVis: Devices: Small & Large — UX
- VIS (Room IV): CG&A Session 1

12:40 PM LUNCH
- Restructuring IEEE VIS for the Future (1:00–2:20 PM) @ *Conv 1, Sec C*

2:20 PM – 4:00 PM
- VAST: High Dimensional Data — DS
- INFOVIS: Text & Communication — AP
- SCIVIS: Space and Physics — SS
- VIS (Estrel C): SciVis Contest — SS
- VIS (Room IV): Panel: Perspectives in Color Research for Scientific Visualization

4:00 PM BREAK

4:20 PM — AP Fast Forward (Thu & Fri Sessions) (4:20–5:20 PM) @ *Estrel A+B*

AP Posters + Networking + Hiring events (5:20–7:00 PM) @ *Foyers*

7:00 PM – 9:00 PM — VIS Dinner Banquet (7:00–9:00 PM) @ *Conv 1, Sec C+D*

THURSDAY, 25 OCTOBER

Tracks: VAST (Conv 1, Sec C) · INFOVIS (Conv 1, Sec D) · SCIVIS (Estrel A+B) · VIS (Estrel C) · VIS (Room IV)

9:00 AM – 10:40 AM
- VAST: Security, Privacy, and Anomaly
- INFOVIS: Immersive Analytics — UX
- SCIVIS: Tensors
- VIS (Estrel C): VAST: Interactive Analytics and Design — DS
- VIS (Room IV): VisAP: Arts Program Session 2: Paths and Memories

10:40 AM BREAK

11:00 AM – 12:40 PM
- VAST: Deep Learning — DS
- INFOVIS: Design & Storytelling — DJ
- SCIVIS: Scalable Techniques — SS
- VIS (Estrel C): InfoVis: Interaction — UX
- VIS (Room IV): CG&A Session 2 — SS BA

12:40 PM LUNCH
- VIS 2019 Kick-off Meeting (1:00–2:20 PM) @ *Estrel C*

2:20 PM – 4:00 PM
- VAST: Graph and Image
- INFOVIS: Perception & Cognition 1
- SCIVIS: Topology, Geometry, and Precision
- VIS (Estrel C): SciVis Short: Visual Abstractions, Perceptual Study and Immersive Vis — SS UX
- VIS (Room IV): Panel: Meet the Founders: How to Start and Sustain a Business in the Vis Space — AP

4:00 PM BREAK

4:20 PM – 6:00 PM
- VAST: Explainable ML — DS
- INFOVIS: Perception & Cognition 2
- SCIVIS: Interaction and Multivariate Data — UX
- VIS (Estrel C): SciVis Short: Flow, Astrophysics, and Computationally Intensive Data Vis

FRIDAY, 26 OCTOBER

Tracks: VAST (Conv 1, Sec C) · INFOVIS (Conv 1, Sec D) · SCIVIS (Estrel C)

9:00 AM – 10:40 AM
- VAST: Event, Sequence, and ML — DS
- INFOVIS: Uncertainty & Error
- SCIVIS: Time-varying Data — SS

10:40 AM BREAK

11:00 AM
- AP **VIS Capstone** (11:00 AM–12:00 PM)
  Can I Believe What I See?—Information-Theoretic Algorithm Validation
  Joachim M. Buhmann, *ETH Zurich*
- AP VIS Closing (12:00–12:30 PM) @ *Conv 1, Sec C*

VisInPractice
Practitioner Guide

Use this key to attend recommended sessions for Practitioners. See the full practitioner guide at http://visinpractice.org.

- AP All Practitioners
- DS Data Science and Machine Learning
- BA Biomedical Applications
- DJ Data Journalism & News Graphics
- UX User Experience, Human Factors, and Interaction Design
- SS Simulation Science

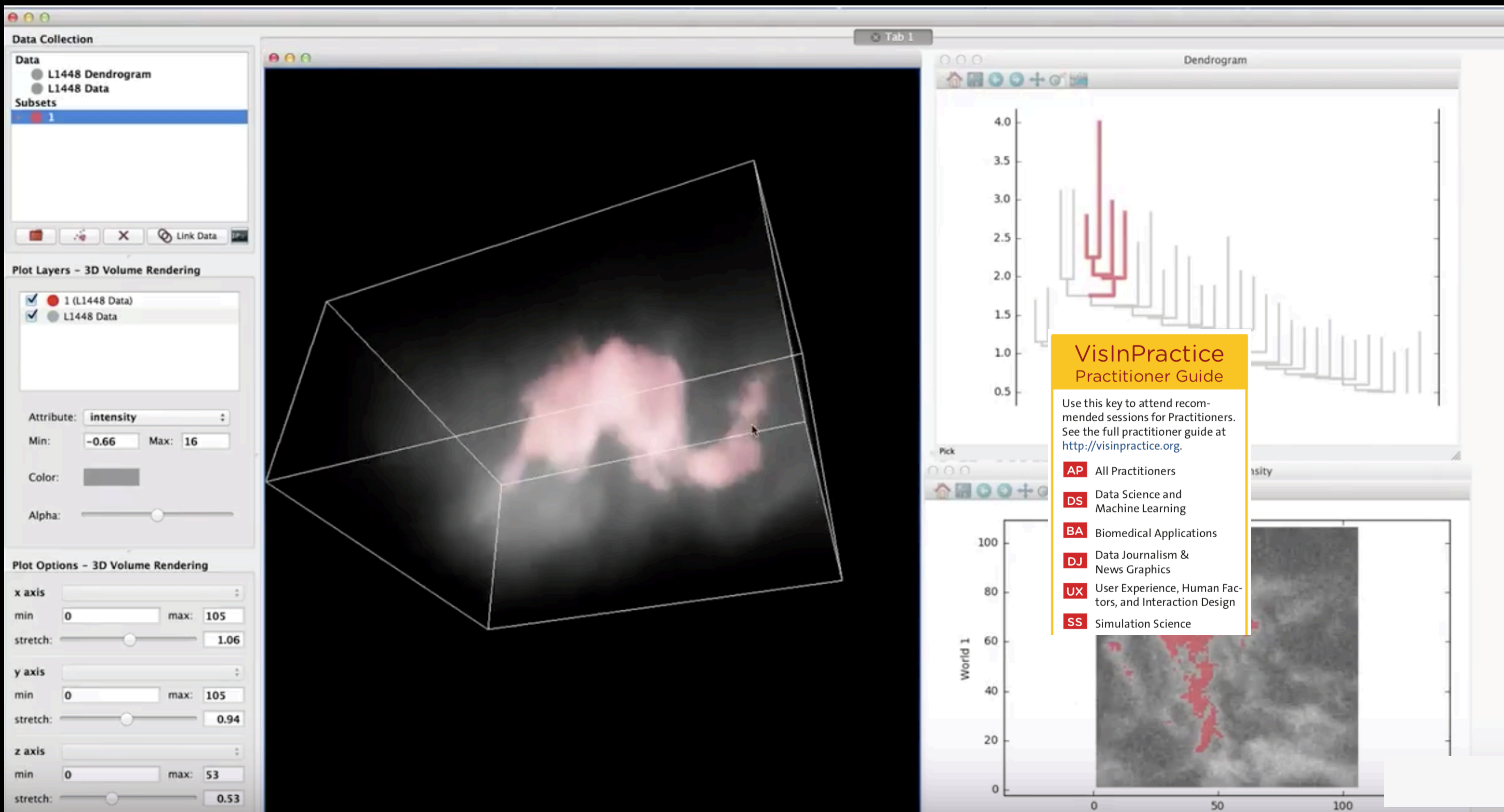
Posters
Located in Foyer 1 & Foyer Estrel Hall
Sunday—Thursday, 9:00 AM–6:00 PM

Exhibitions
Located in Foyer 3
Tuesday—Thursday, 10:40 AM–6:00 PM

VIS Arts Program
Located in Room Paris
Opening, Tuesday 7:00 PM
Exhibition, Wednesday—Thursday, 9:00 AM–6:00 PM

“Information Visualization” “VAST” “Scientific Visualization”

Visualization in Computer Science



Visualization in Science

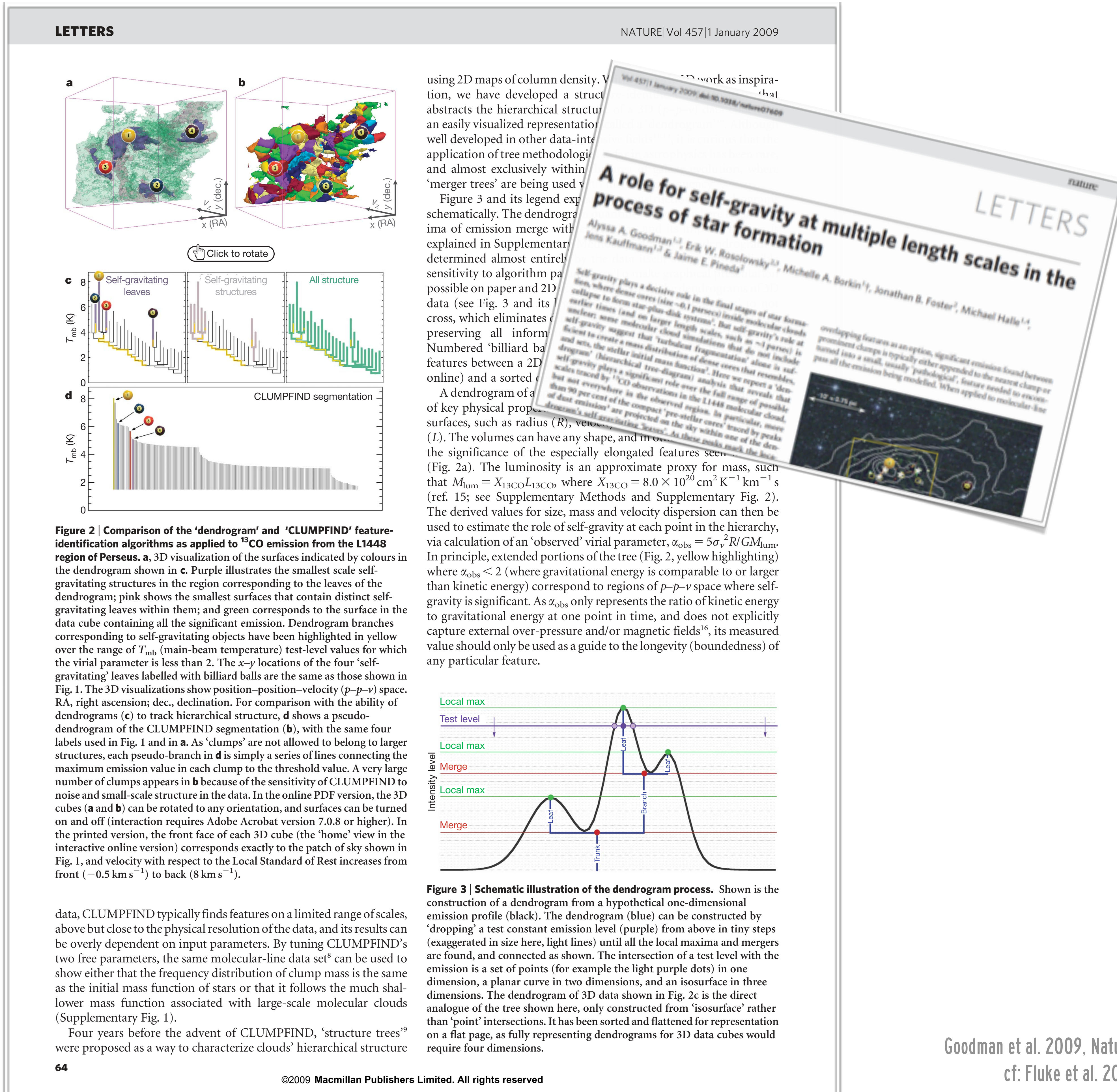


Figure 2 | Comparison of the 'dendrogram' and 'CLUMPFIND' feature-identification algorithms as applied to ^{13}CO emission from the L1448 region of Perseus. a, 3D visualization of the surfaces indicated by colours in the dendrogram shown in **c**. Purple illustrates the smallest scale self-gravitating structures in the region corresponding to the leaves of the dendrogram; pink shows the smallest surfaces that contain distinct self-gravitating 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_{mb} (main-beam temperature) test-level values for which the virial parameter is less than 2. The *x*–*y* locations of the four 'self-gravitating' leaves labelled with billiard balls are the same as those shown in Fig. 1. The 3D visualizations show position–position–velocity (*p*–*p*–*v*) space. RA, right ascension; dec., declination. For comparison with the ability of dendrograms (**c**) to track hierarchical structure, **d** shows a pseudo-dendrogram 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 km s^{-1}) to back (8 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[8] 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'[9] were proposed as a way to characterize clouds' hierarchical structure

using 2D maps of column density. V… …D work as inspiration, we have developed a struct… …that abstracts the hierarchical structur… an easily visualized representation… well developed in other data-inte… application of tree methodologi… and almost exclusively within… 'merger trees' are being used …

Figure 3 and its legend exp… schematically. The dendrogra… ima of emission merge with… explained in Supplementary… determined almost entirely… sensitivity to algorithm pa… possible on paper and 2D… data (see Fig. 3 and its l… cross, which eliminates… preserving all inform… Numbered 'billiard ba… features between a 2D… online) and a sorted…

A dendrogram of a… of key physical prope… surfaces, such as radius (R), veloc… (L). The volumes can have any shape, and in ou… the significance of the especially elongated features seen… (Fig. 2a). The luminosity is an approximate proxy for mass, such that $M_{lum} = X_{13CO}L_{13CO}$, where $X_{13CO} = 8.0 \times 10^{20}$ cm^2 K^{-1} km^{-1} s (ref. 15; see Supplementary Methods and Supplementary Fig. 2). The derived values for size, mass and velocity dispersion can then be used to estimate the role of self-gravity at each point in the hierarchy, via calculation of an 'observed' virial parameter, $\alpha_{obs} = 5\sigma_v^2 R/GM_{lum}$. In principle, extended portions of the tree (Fig. 2, yellow highlighting) where $\alpha_{obs} < 2$ (where gravitational energy is comparable to or larger than kinetic energy) correspond to regions of *p*–*p*–*v* space where self-gravity is significant. As α_{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[16], 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.

Vol 457|1 January 2009|doi:10.1038/nature07609

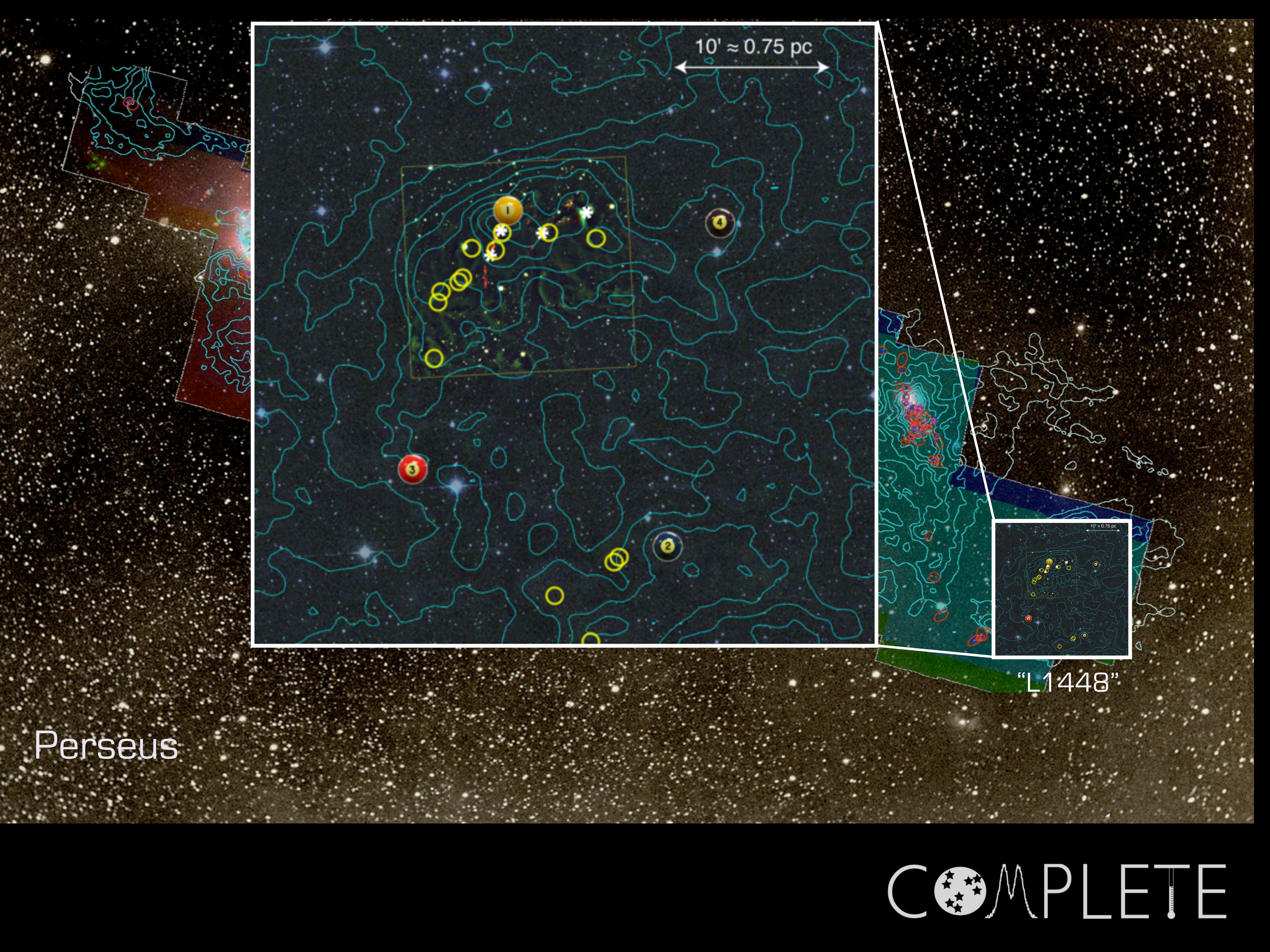
nature

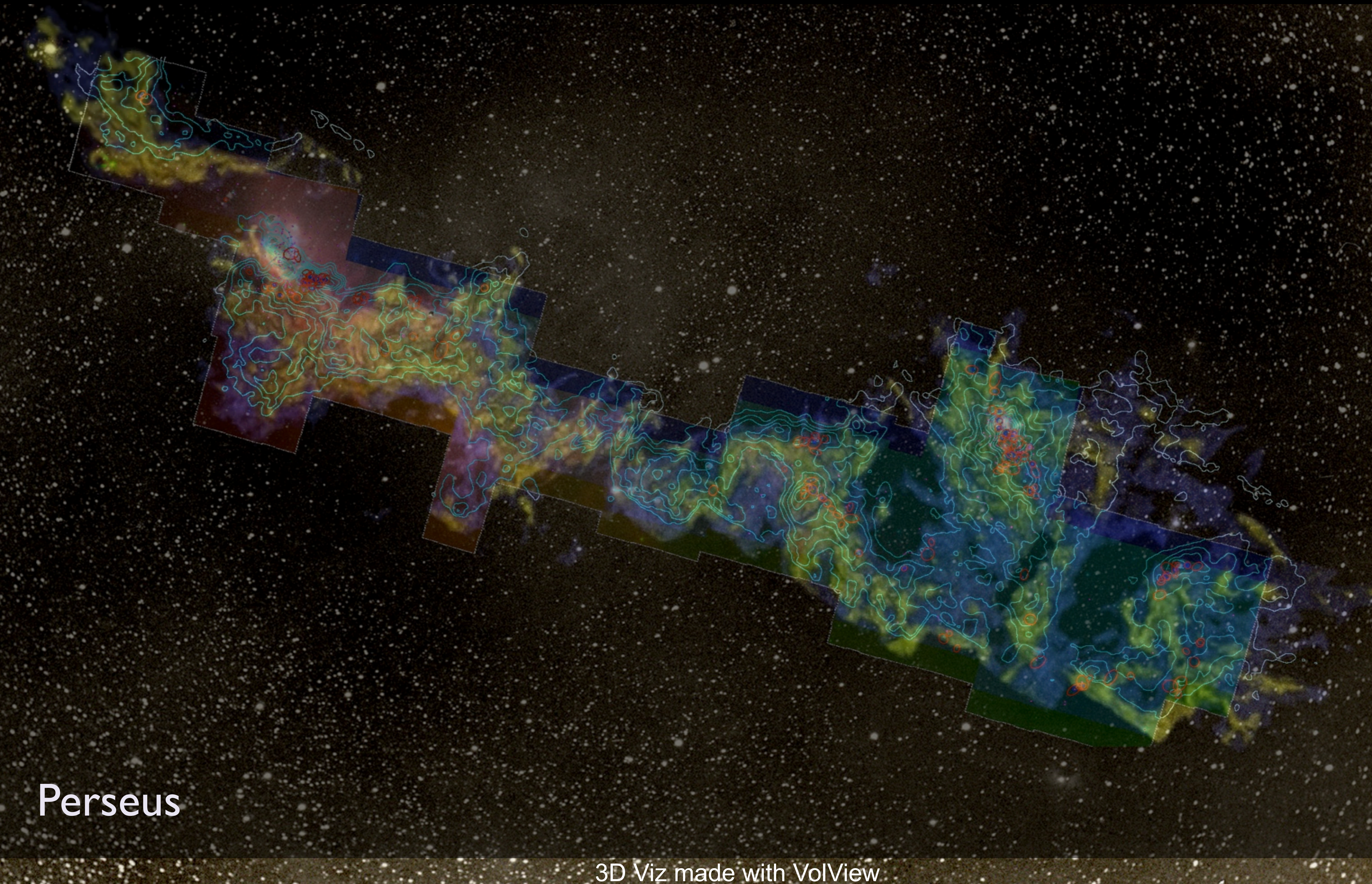
LETTERS

A role for self-gravity at multiple length scales in the process of star formation

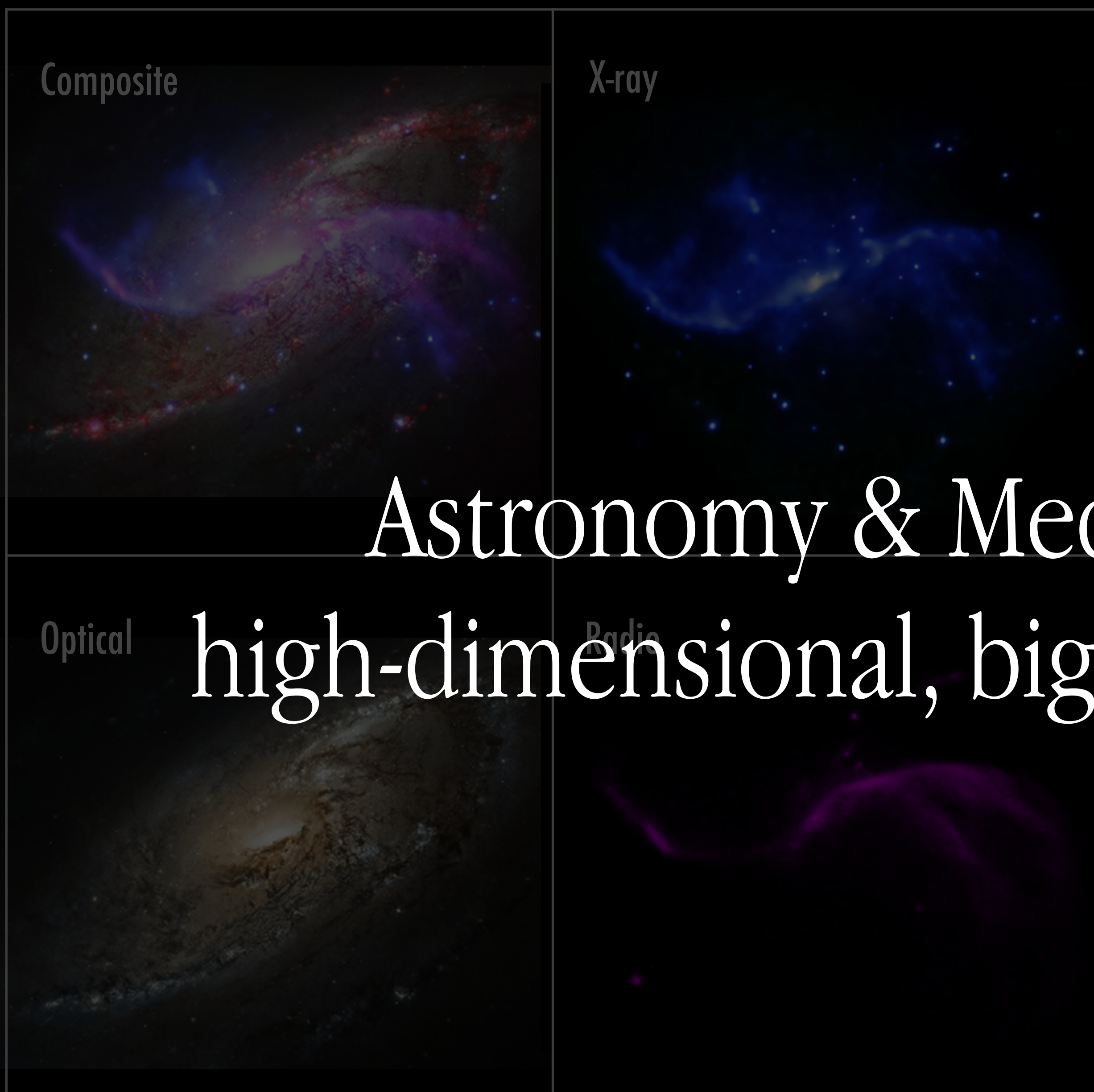
Alyssa A. Goodman[1,2], Erik W. Rosolowsky[2,3], Michelle A. Borkin[1], Jonathan B. Foster[2], Michael Halle[1,4], Jens Kauffmann[1,2] & Jaime E. Pineda[2]

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


10' ≈ 0.75 pc
"L1448"
Perseus
COMPLETE


Perseus
3D Viz made with VolView

Why AstronomicalMedicine?



Astronomy & Medicine both rely on high-dimensional, big, wide, data for insight.

chandra.harvard.edu/photo/2014/m106/

Chang, et al. 2011, brain.oxfordjournals.org/content/134/12/3632

Visualization in Computer Science


LETTERS

A role for self-gravity at multiple length scales in the process of star formation

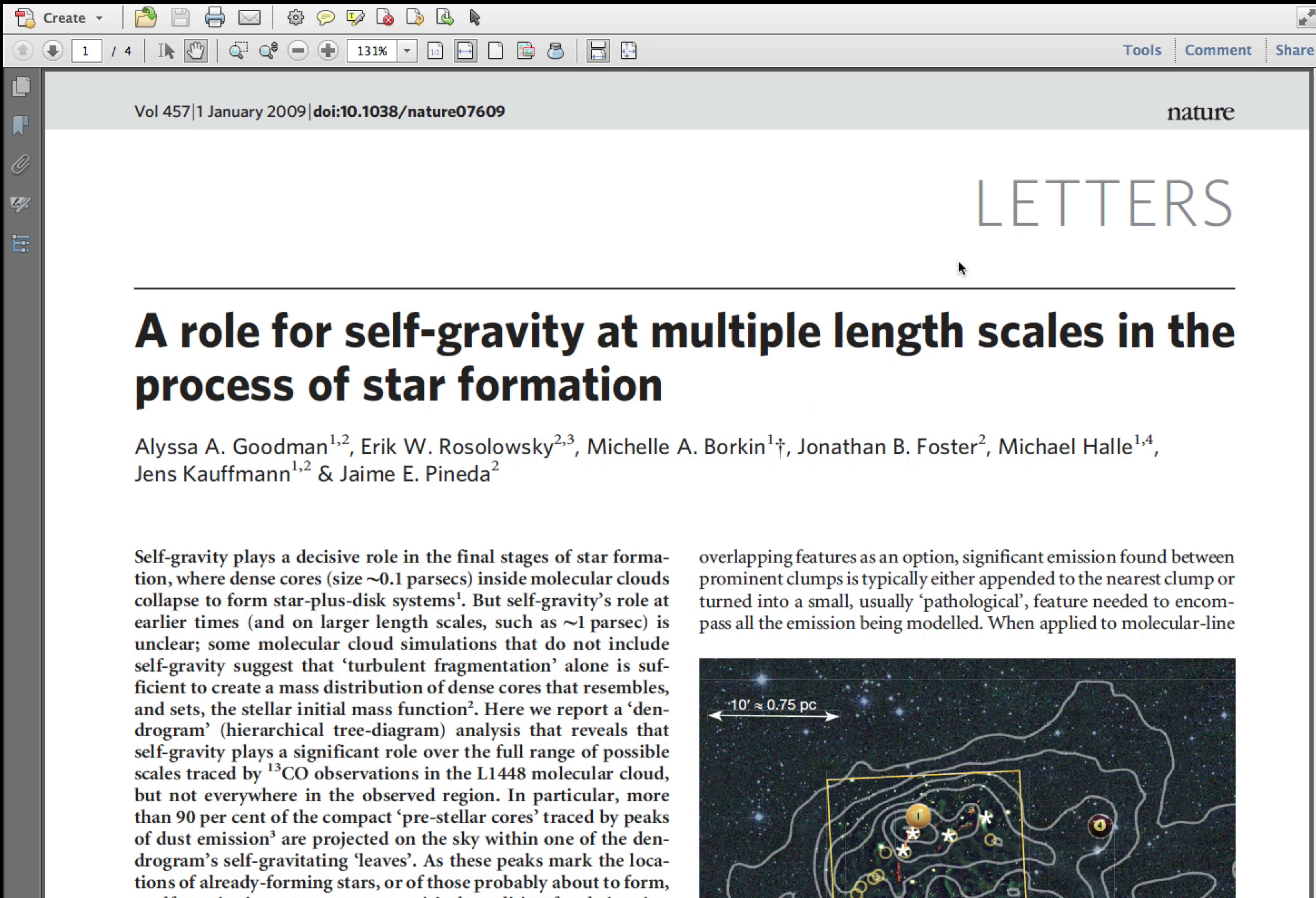
“VAST”

“Scientific Visualization”

“Information Visualization”

2009: “3D PDF”

“VAST”



Vol 457 | 1 January 2009 | doi:10.1038/nature07609 nature

LETTERS

A role for self-gravity at multiple length scales in the process of star formation

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

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 systems[1]. 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 function[2]. 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 ^{13}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 emission[3] 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,

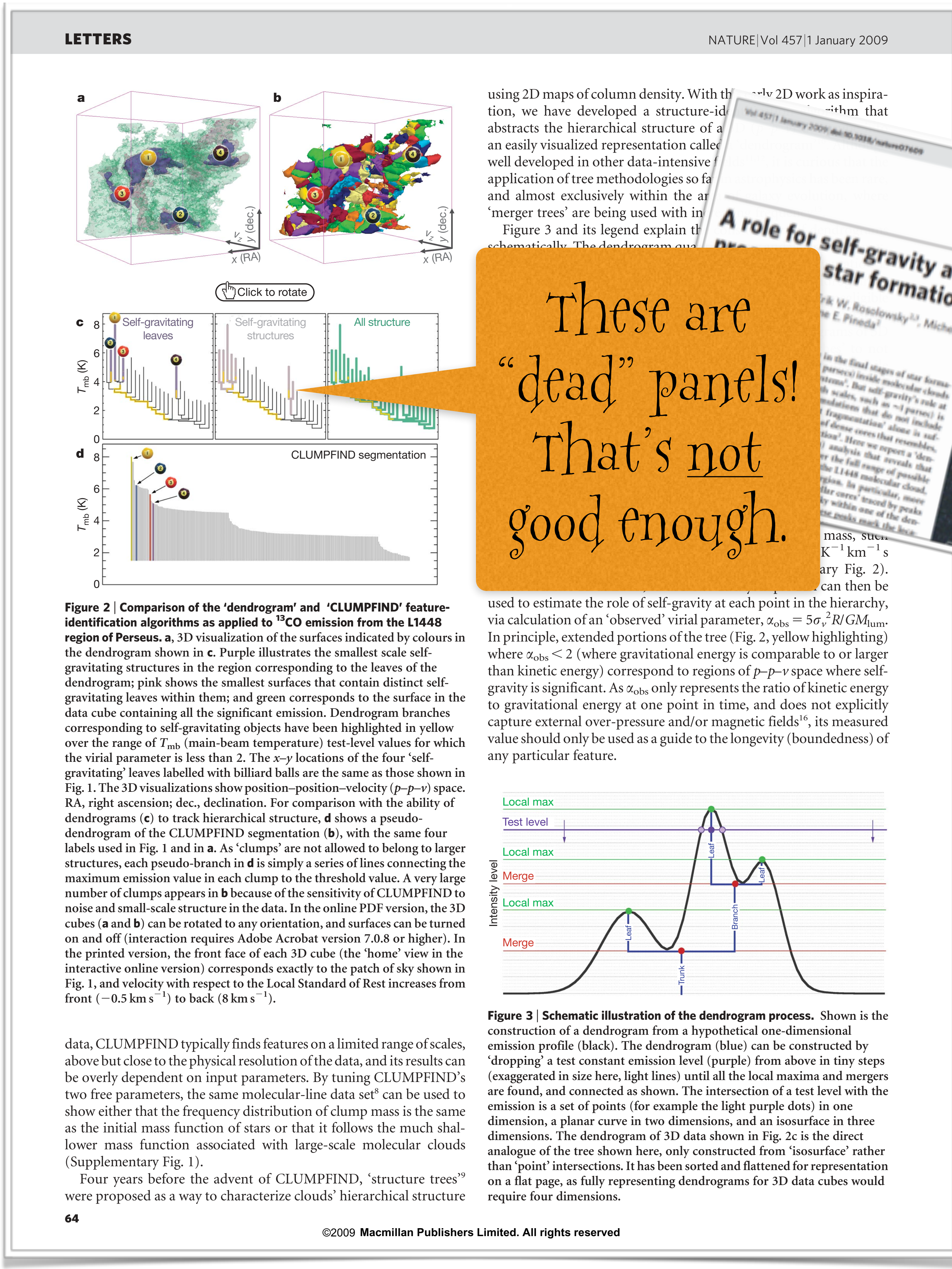
overlapping features as an option, significant emission found between prominent clumps is typically either appended to the nearest clump or turned into a small, usually ‘pathological’, feature needed to encompass all the emission being modelled. When applied to molecular-line

“Visual analytics is the science of analytical reasoning supported by highly interactive visual interfaces”

2009

3D PDF

High-Dimensional data in a "Paper" on its way to the Future



a b

x (RA) y (dec.) v

Click to rotate

c Self-gravitating leaves | Self-gravitating structures | All structure

d CLUMPFIND segmentation

T_{mb} (K)

Figure 2 | Comparison of the 'dendrogram' and 'CLUMPFIND' feature-identification algorithms as applied to ^{13}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 self-gravitating structures in the region corresponding to the leaves of the dendrogram; pink shows the smallest surfaces that contain distinct self-gravitating 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_{mb} (main-beam temperature) test-level values for which the virial parameter is less than 2. The *x*–*y* locations of the four 'self-gravitating' leaves labelled with billiard balls are the same as those shown in Fig. 1. The 3D visualizations show position–position–velocity (*p*–*p*–*v*) space. RA, right ascension; dec., declination. For comparison with the ability of dendrograms (**c**) to track hierarchical structure, **d** shows a pseudo-dendrogram 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 ($-0.5\ km\ s^{-1}$) to back ($8\ km\ s^{-1}$).

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[8] 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'[9] were proposed as a way to characterize clouds' hierarchical structure using 2D maps of column density. With th... ...rly 2D work as inspiration, we have developed a structure-id... ...rithm that abstracts the hierarchical structure of a... an easily visualized representation calle... well developed in other data-intensive f... application of tree methodologies so fa... and almost exclusively within the ar... 'merger trees' are being used with in...

Figure 3 and its legend explain th... schematically. The dendrogram qua...

... mass, such ... $K^{-1}\ km^{-1}\ s$ ... ary Fig. 2). ... can then be used to estimate the role of self-gravity at each point in the hierarchy, via calculation of an 'observed' virial parameter, $\alpha_{obs} = 5\sigma_v^2 R/GM_{lum}$. In principle, extended portions of the tree (Fig. 2, yellow highlighting) where $\alpha_{obs} < 2$ (where gravitational energy is comparable to or larger than kinetic energy) correspond to regions of *p*–*p*–*v* space where self-gravity is significant. As α_{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[16], its measured value should only be used as a guide to the longevity (boundedness) of any particular feature.

Local max
Test level
Local max
Merge
Local max
Merge
Intensity level
Leaf
Leaf
Leaf
Branch
Trunk

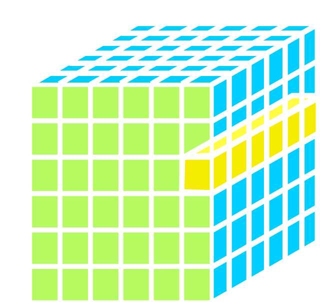
Figure 3 | Schematic illustration of the dendrogram process. Shown is the construction of a dendrogram from a hypothetical one-dimensional emission profile (black). The dendrogram (blue) can be constructed by 'dropping' a test constant emission level (purple) from above in tiny steps (exaggerated in 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.

nature LETTERS

Vol 457|1 January 2009|doi:10.1038/nature07609

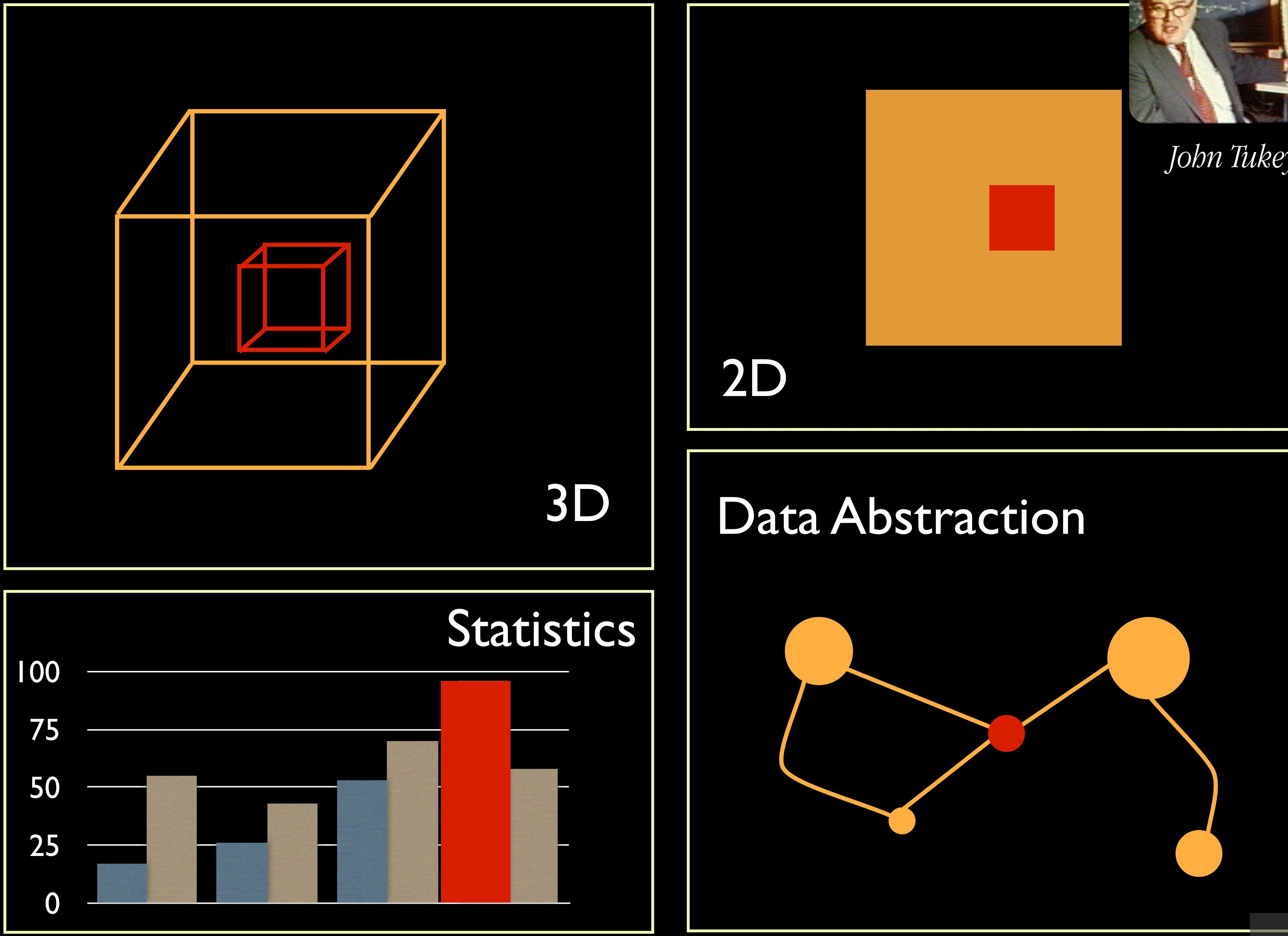
A role for self-gravity at multiple length scales in the process of star formation

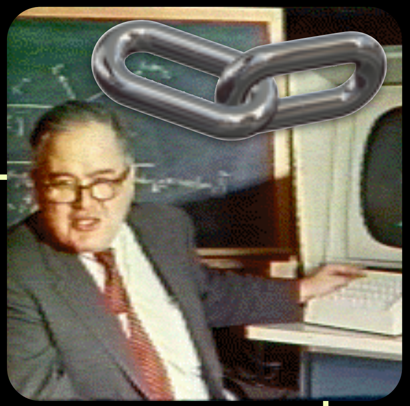
...rk W. Rosolowsky, Michelle A. Borkin, Jonathan B. Foster, Michael Halle, ...e E. Pineda



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

Linked Views of High-dimensional Data

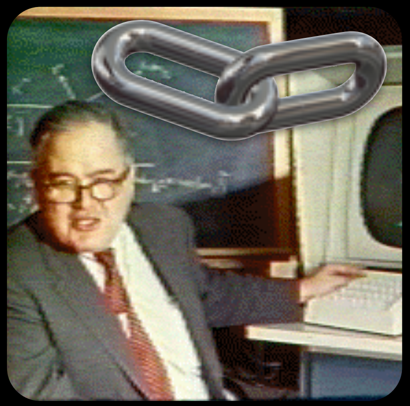




John Tukey

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

John Tukey's Legacy



PRIM-9

PRIM-H

DataDesk®

XGobi

GGobi

RGGobi

glue

Spotfire®

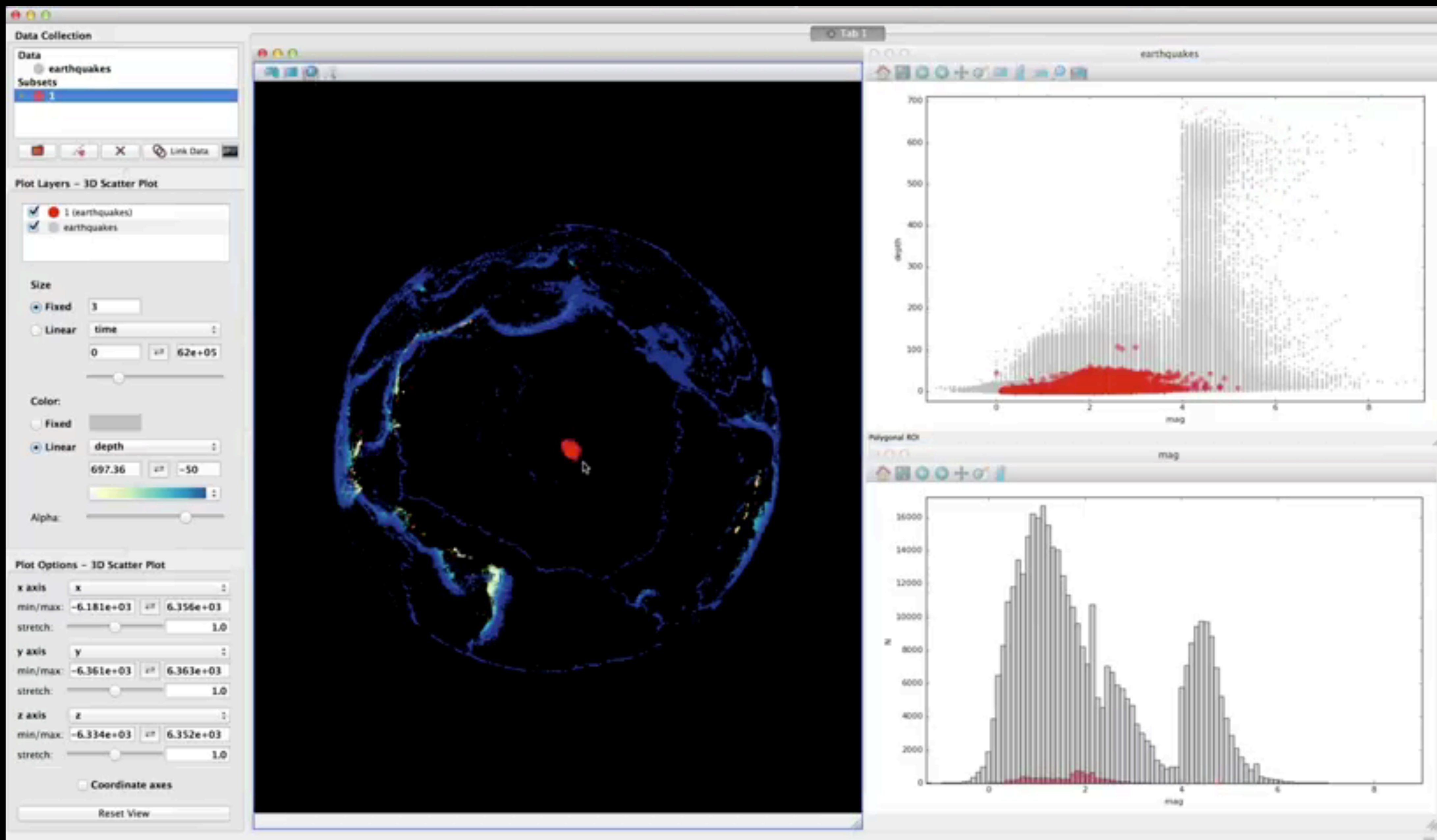
Polaris

tableau

1970 1980 1990 2000 2010

Linked Views of High-dimensional Data (in Python)

glue

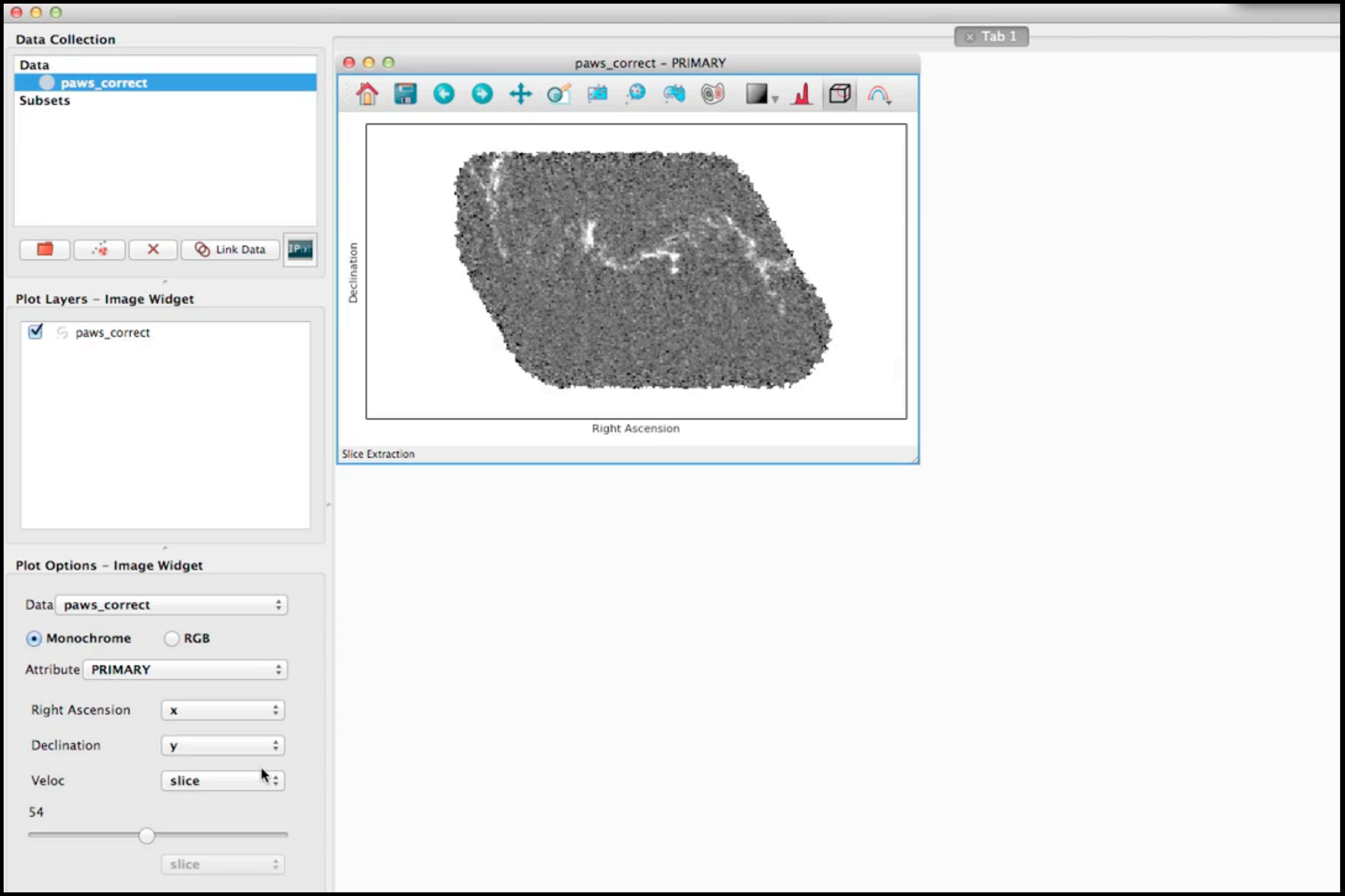


video by Tom Robitaille, lead glue developer

glue created by: C. Beaumont, M. Borkin, P. Qian, T. Robitaille, and A. Goodman, PI

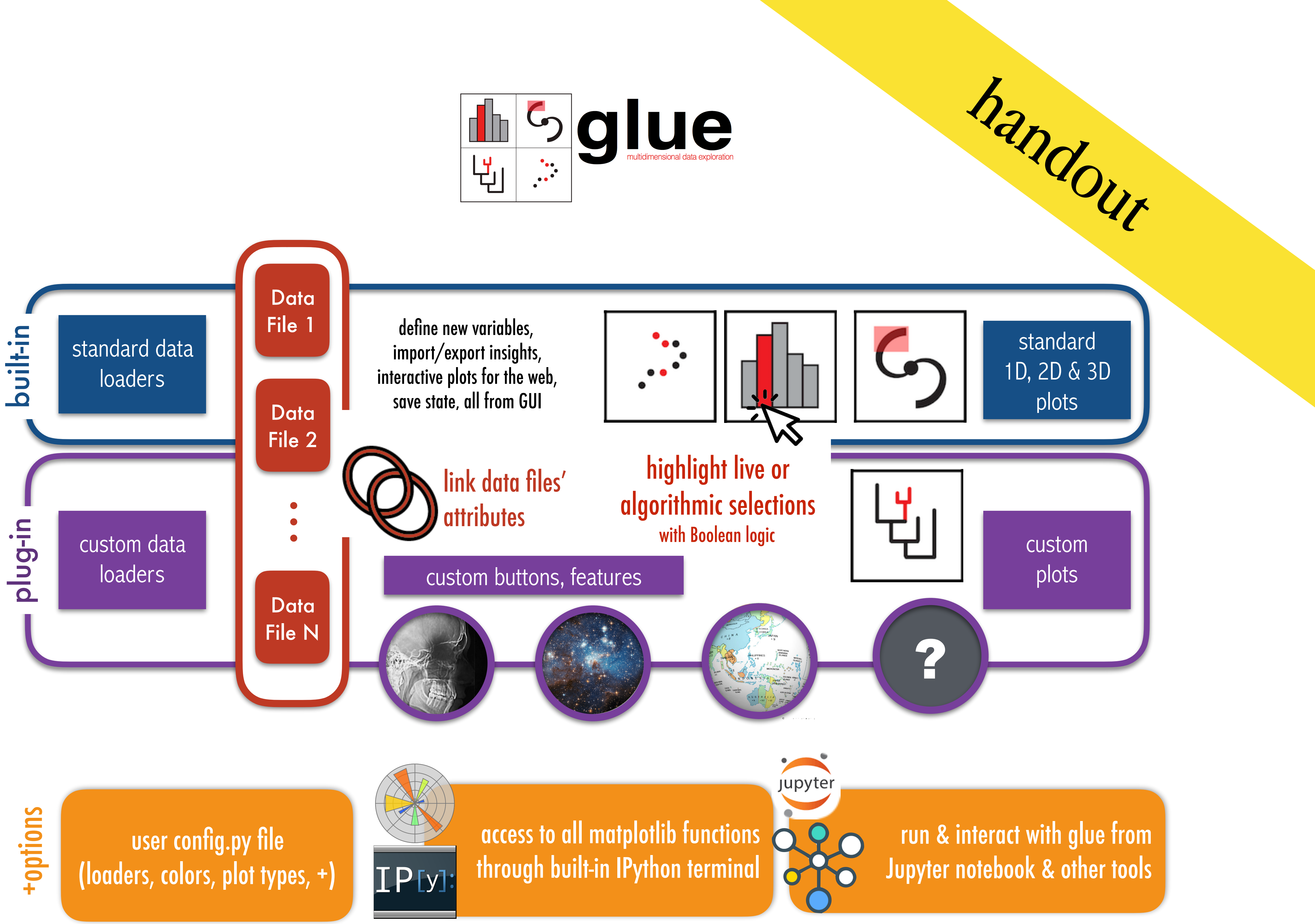
Linked Views of High-dimensional Data (in Python)

glue



video by Chris Beaumont, glue developer

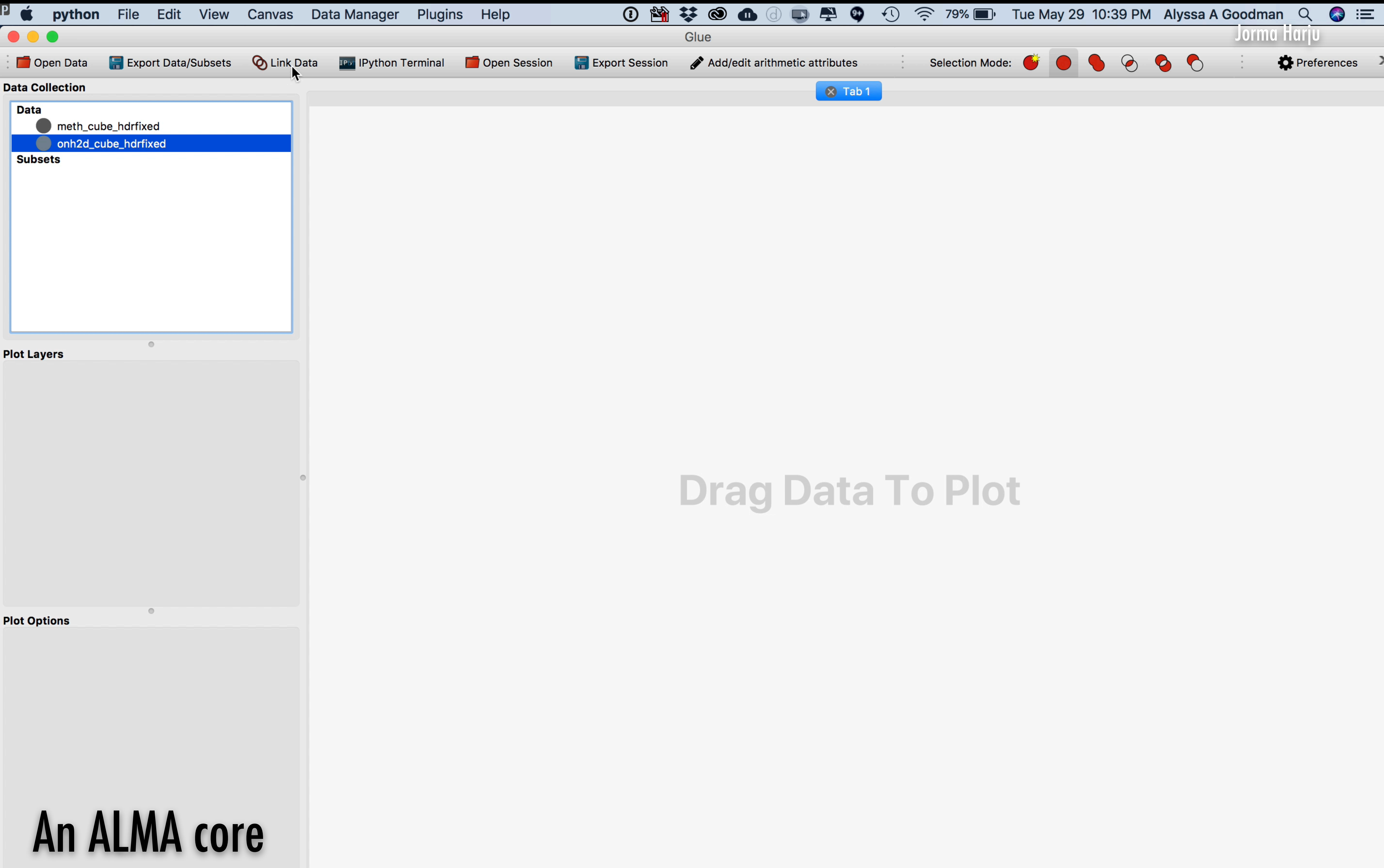
glue created by: C. Beaumont, M. Borkin, P. Qian, T. Robitaille, and A. Goodman, PI



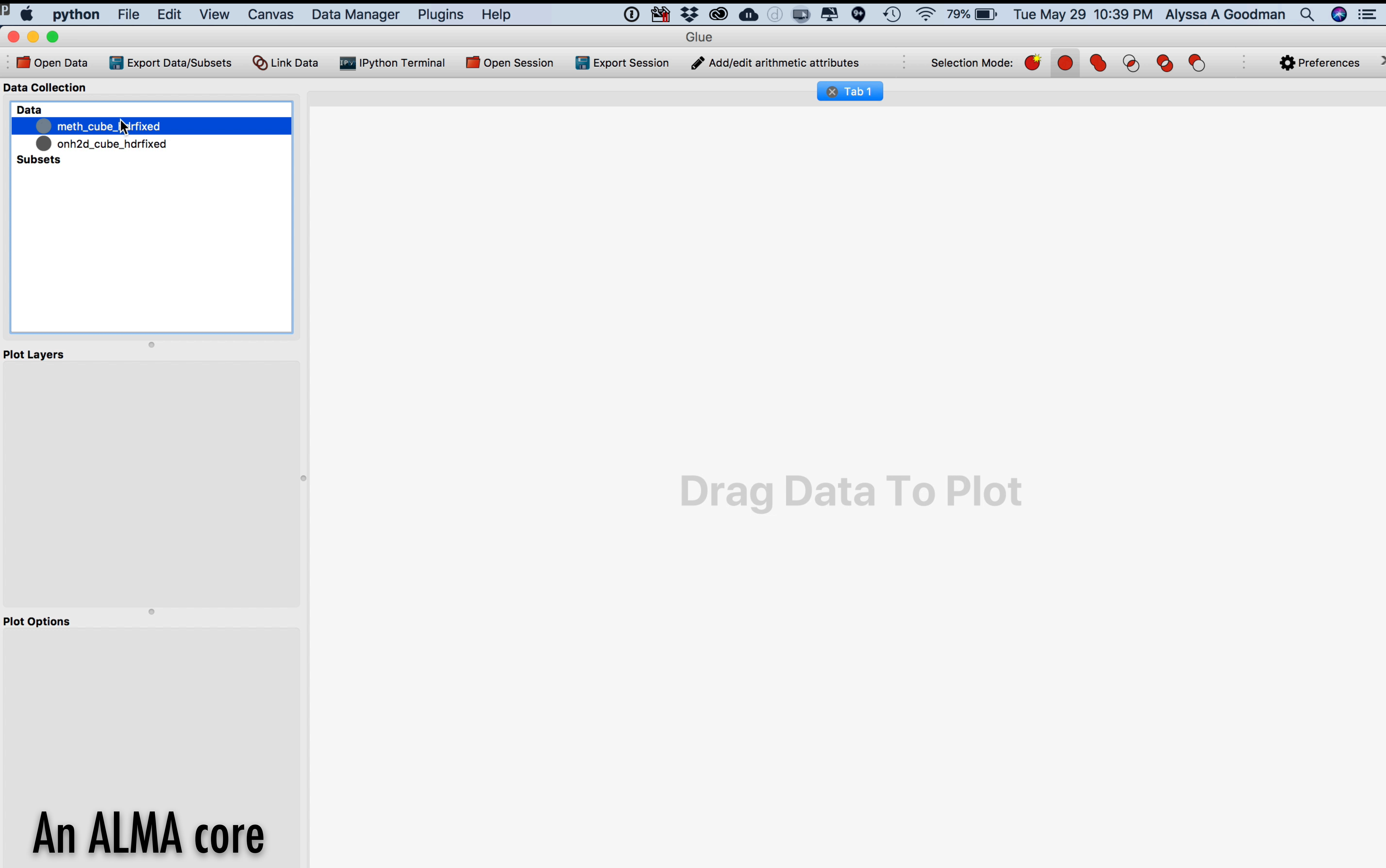
glueviz.org

[demo]
[ALMA]
[publishing]

No merging of data sets—just glue them.

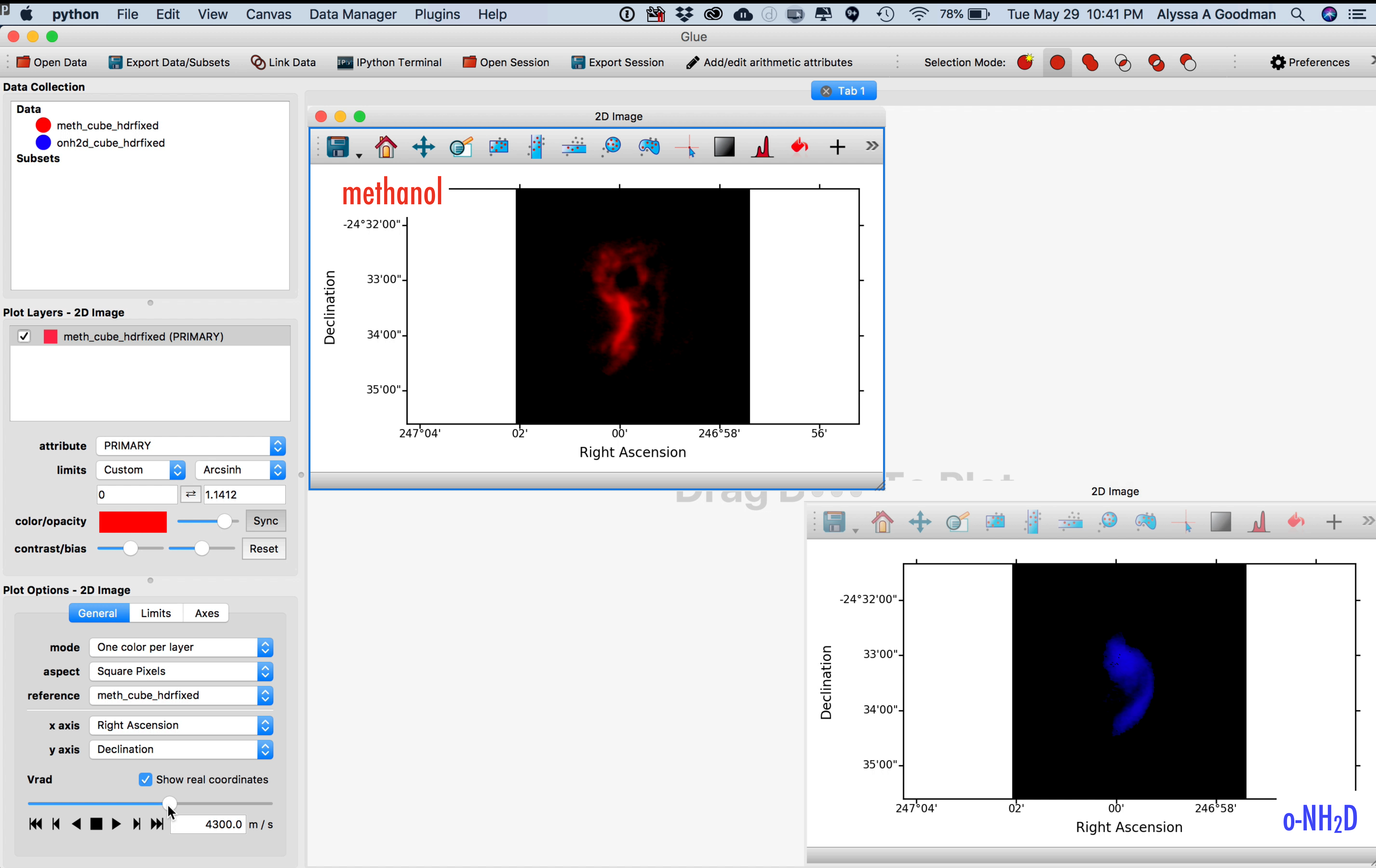


Just drag to visualize, e.g. series of 2D "channel maps."

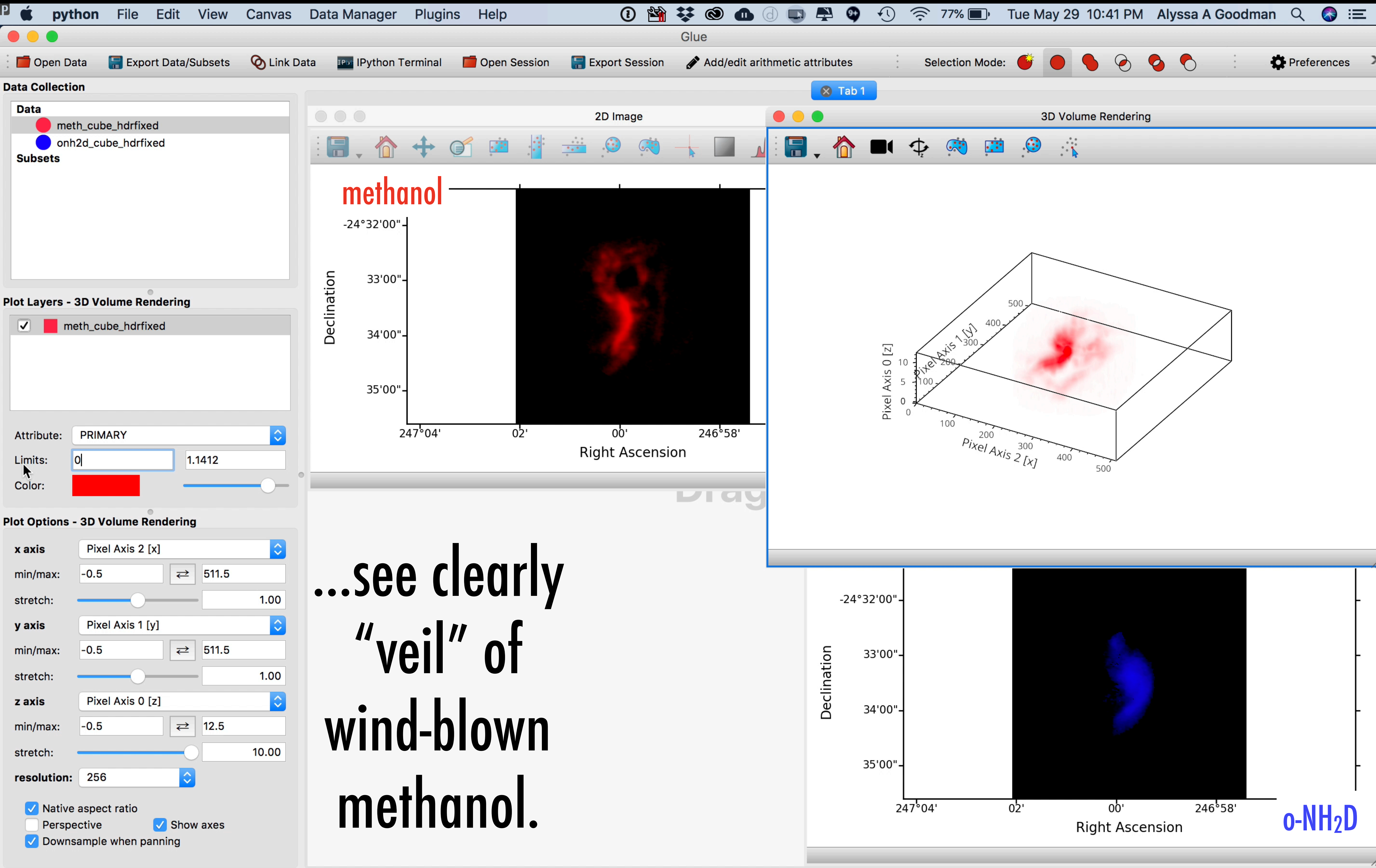


An ALMA core

Adjust so each tracer is a different color.



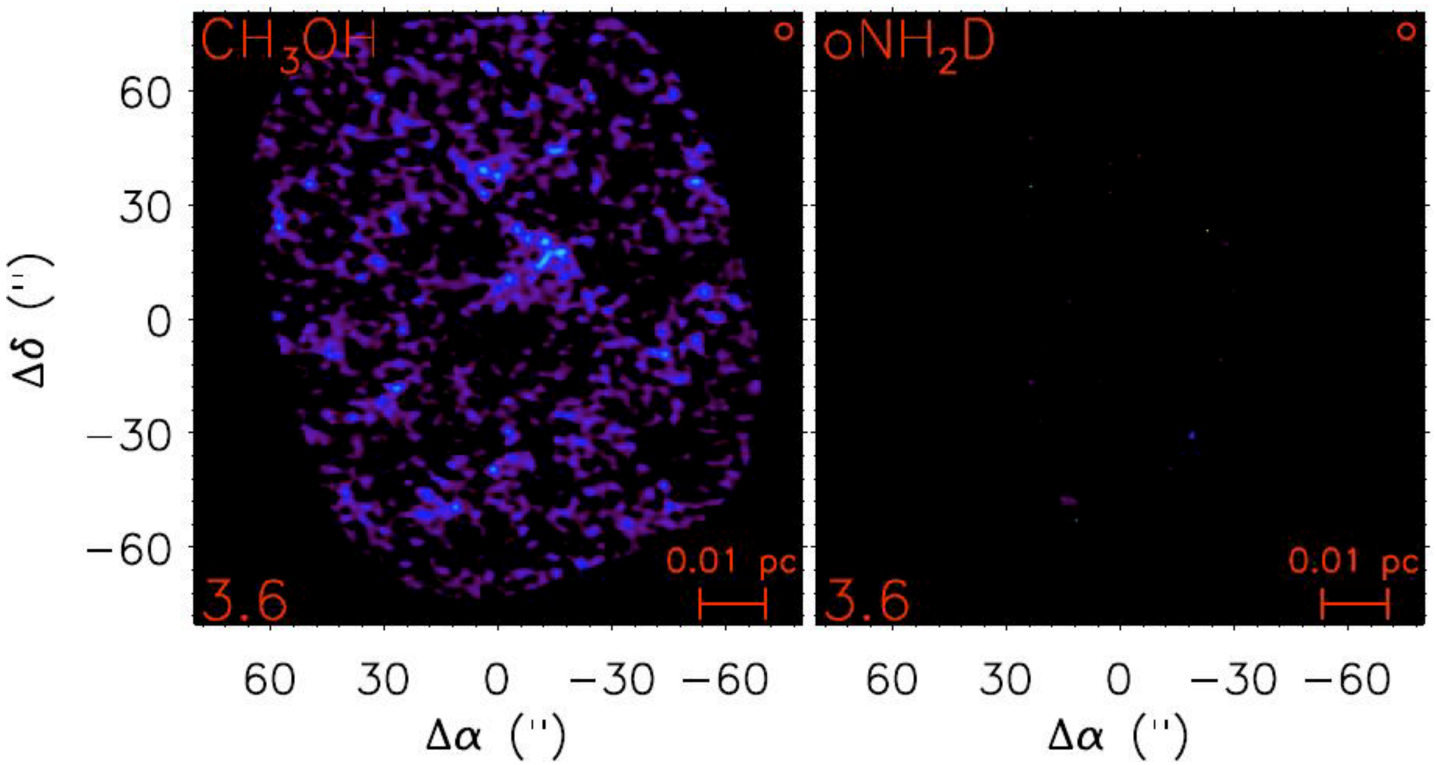
Create 3D views...

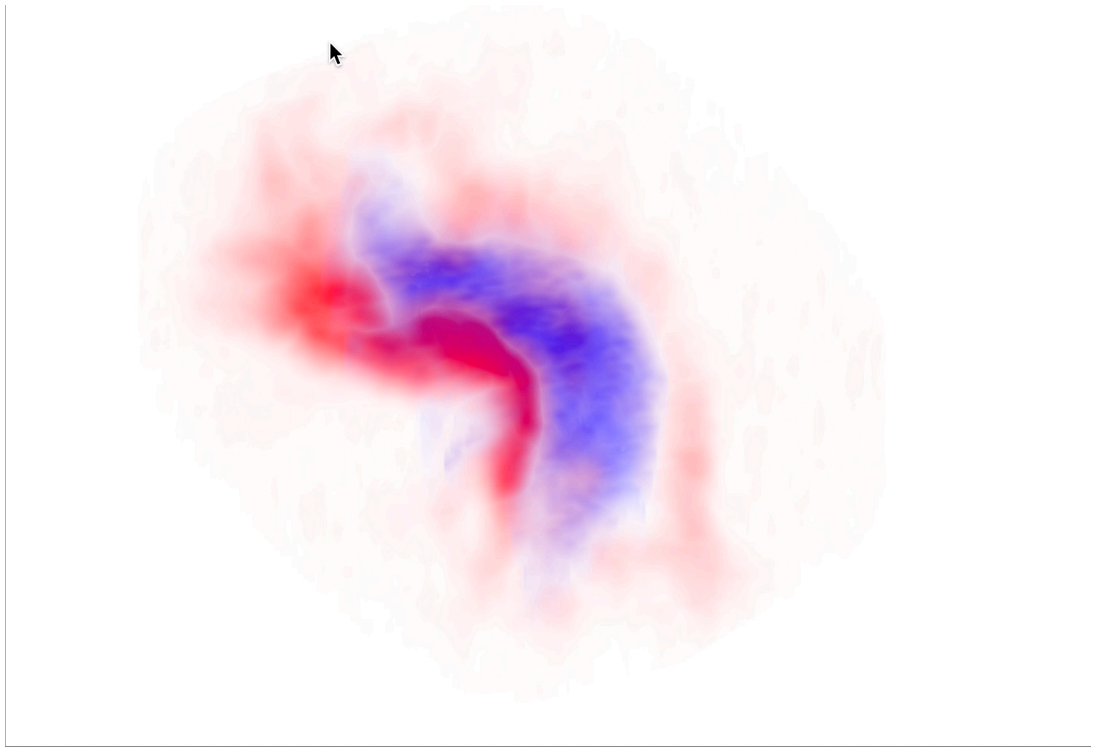


...see clearly "veil" of wind-blown methanol.

Traditional Rainbow Channel maps

glue





2015: The "Paper" of the Future

Au thorea FEATURED ARTICLES ABOUT PLANS BLOG FEEDBACK HELP ALYSSA GOODMAN

PUBLIC ROUGH DRAFT Index Settings Fork Quickedit Word Count 42 Comments Export Unfollow

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 Re-arrange authors

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

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.

Konrad Hinsen 3 days ago · Public
Many good suggestions, but if the goal is "long-lasting rich records of scientific discourse", a more careful and critical attitude towards electronic artifacts is appropriate. I do see it concerning videos, but not a word on the much more critical situation in software. Archiving source code is not sufficient: all the dependencies, plus the complete build environment, would have to be conserved as well to make things work a few years from now. An "executable figure" in the form of an IPython notebook wil...
more

Merce Crosas 3 days ago · Public
Konrad, good points; this has been a concern for the community working on reproducibility. Regarding data repositories, Dataverse handles long-term preservation and access of data files in the following way: 1) for some data files that the repository recognizes (such as R Data, SPSS, STATA), which depend on a statistical package, the system converts them into a preservation format (such as a tab/CSV format). Even though the original format is also saved and can be accessed, the new preservation format gua...
more

Konrad Hinsen 1 day ago · Public
That sounds good. I hope more repositories will follow the example of Dataverse. Figshare in particular has a very different attitude, encouraging researchers to deposit as much as possible. That's perhaps a good strategy to change habits, but in the long run it could well backfire when people find out in a few years that 90% of those deposits have become useless.

Christine L. Borgman 4 months ago · Private
"publications"

d3po/Authorea: Peek, Price-Whelan, Pepe, Beaumont, Borkin, Newton; PoF: Goodman, Peek; WWT: Wong, Fay et al.; Astrometry.net: Hogg, Lang, Roweis et al.

[demo]

[skip arch]

glue
multidimensional data exploration

plotly

enabled by d3.js (javascript) outputs

d3po

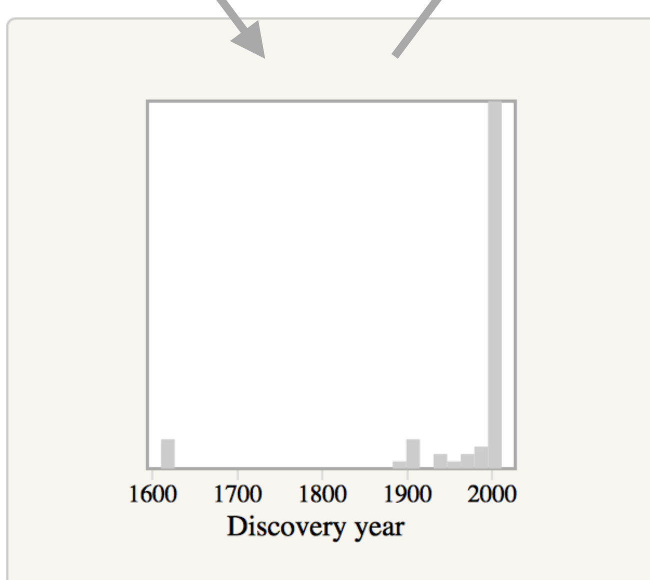
d3po is a project designed to allow an astronomer (or anyone), with no special data visualization skills, to make an interactive, publication-quality figure that has staged builds and linked brushing through scatter plots. Our current version can be previewed at d3po.org, and represents a figure from upcoming work by graduate student Elisabeth Newton. The figure describes how metalicity affects color in cool stars, and represents a nice use case for d3po. Try clicking and dragging in the scatter plots to understand the power of linked brushing in published figures.

Right now we are in search of alpha testers, who have figures that could be made interactive and who are willing to get their hands a little dirty (No javascript skills needed). In future versions, we plan to link to glue to allow the creation of d3po figures interactively. We are also exploring implementation of d3po within presentations and within authorea. Full 1.0 version expected in January 2014.

Installing your own d3po server

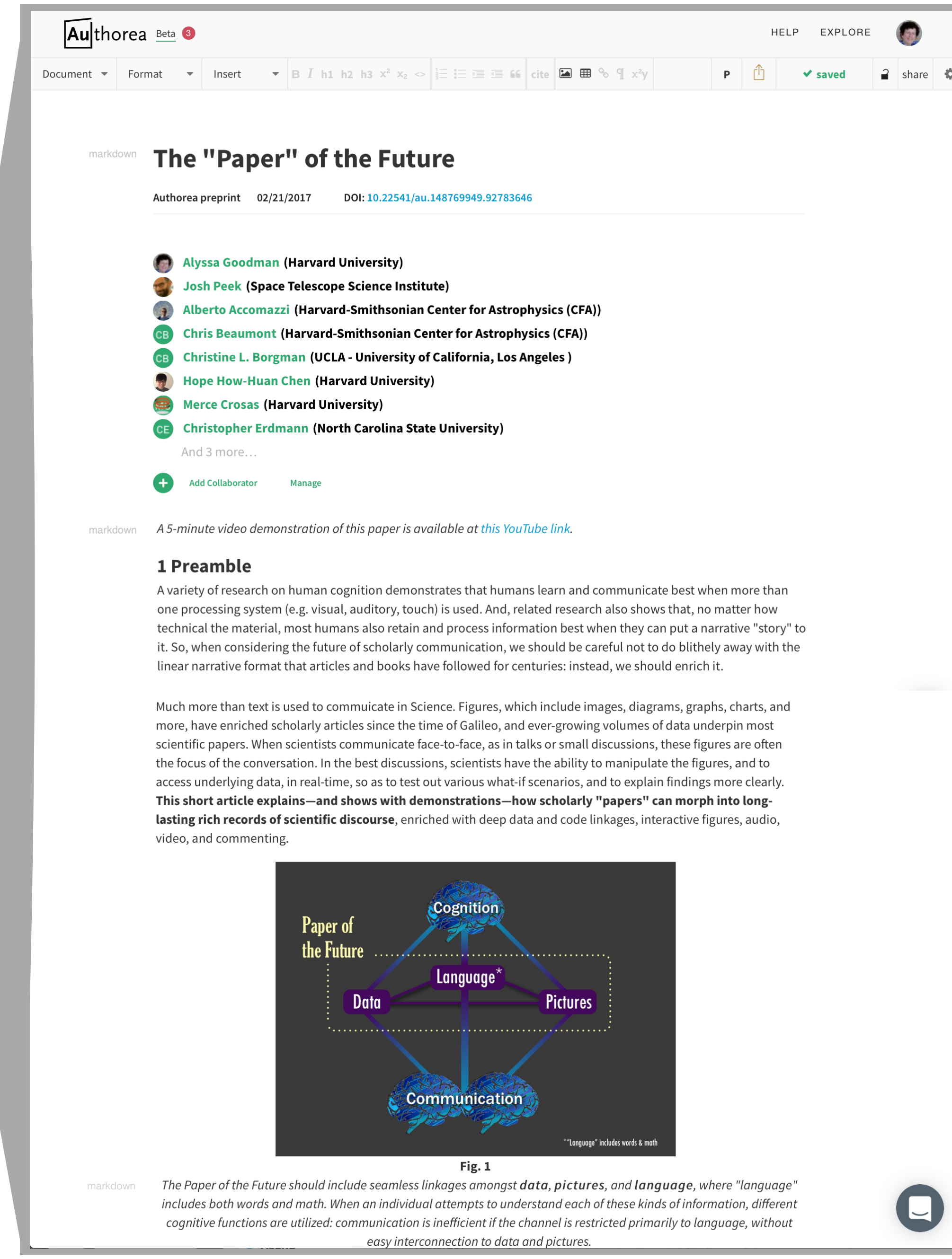
```
git clone git@github.com:adrn/d3po.git
cd d3po
virtualenv --no-site-packages venv
source venv/bin/activate
pip install -r pip-requirements.txt
python run.py
```

Authorea



Four Centuries of Discovery | A Chasm in Mass | Little Siblings | Close Cousins | The Strangers

After Galileo discovered the first four moons of Jupiter, it took nearly three hundred years to discover the next one.



Authorea Beta | HELP | EXPLORE

Document | Format | Insert | saved | share

The "Paper" of the Future

Authorea preprint 02/21/2017 DOI: 10.22541/au.148769949.92783646

Alyssa Goodman (Harvard University)
Josh Peek (Space Telescope Science Institute)
Alberto Accomazzi (Harvard-Smithsonian Center for Astrophysics (CFA))
Chris Beaumont (Harvard-Smithsonian Center for Astrophysics (CFA))
Christine L. Borgman (UCLA - University of California, Los Angeles)
Hope How-Huan Chen (Harvard University)
Merce Crosas (Harvard University)
Christopher Erdmann (North Carolina State University)
And 3 more...

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

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.

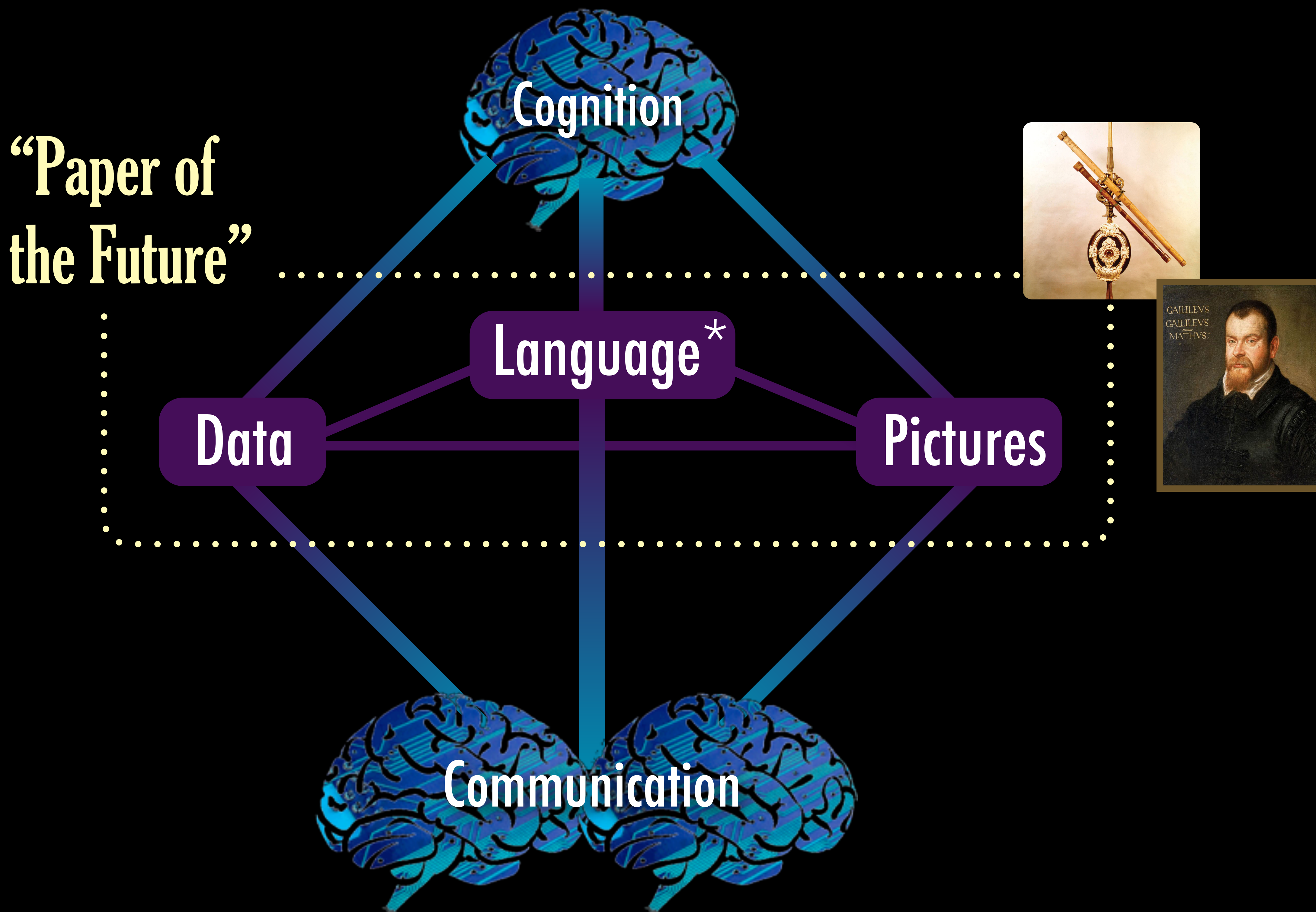
Much more than text is used to commuicate in Science. Figures, which include images, diagrams, graphs, charts, and more, have enriched scholarly articles since the time of Galileo, and ever-growing volumes of data underpin most scientific papers. When scientists communicate face-to-face, as in talks 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.

Fig. 1

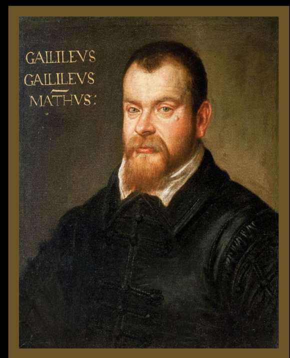
*The Paper of the Future should include seamless linkages amongst **data**, **pictures**, and **language**, where "language" includes both words and math. When an individual attempts to understand each of these kinds of information, different cognitive functions are utilized: communication is inefficient if the channel is restricted primarily to language, without easy interconnection to data and pictures.*

[demo]

Many thanks to Alberto Pepe, Josh Peek, Chris Beaumont, Tom Robitaille, Adrian Price-Whelan, Elizabeth Newton, Michelle Borkin & Matteo Cantiello for making this posible.

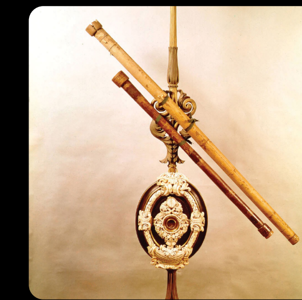


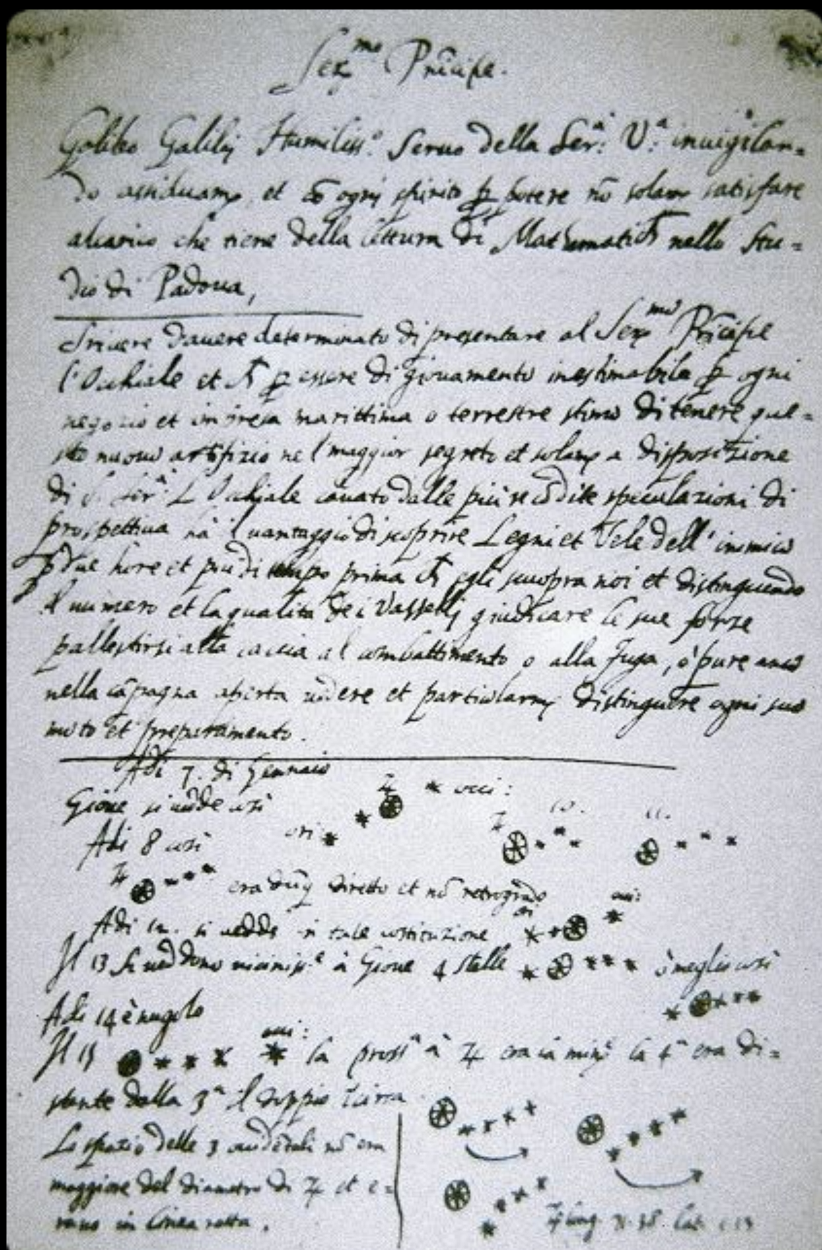
*"Language" includes words & math

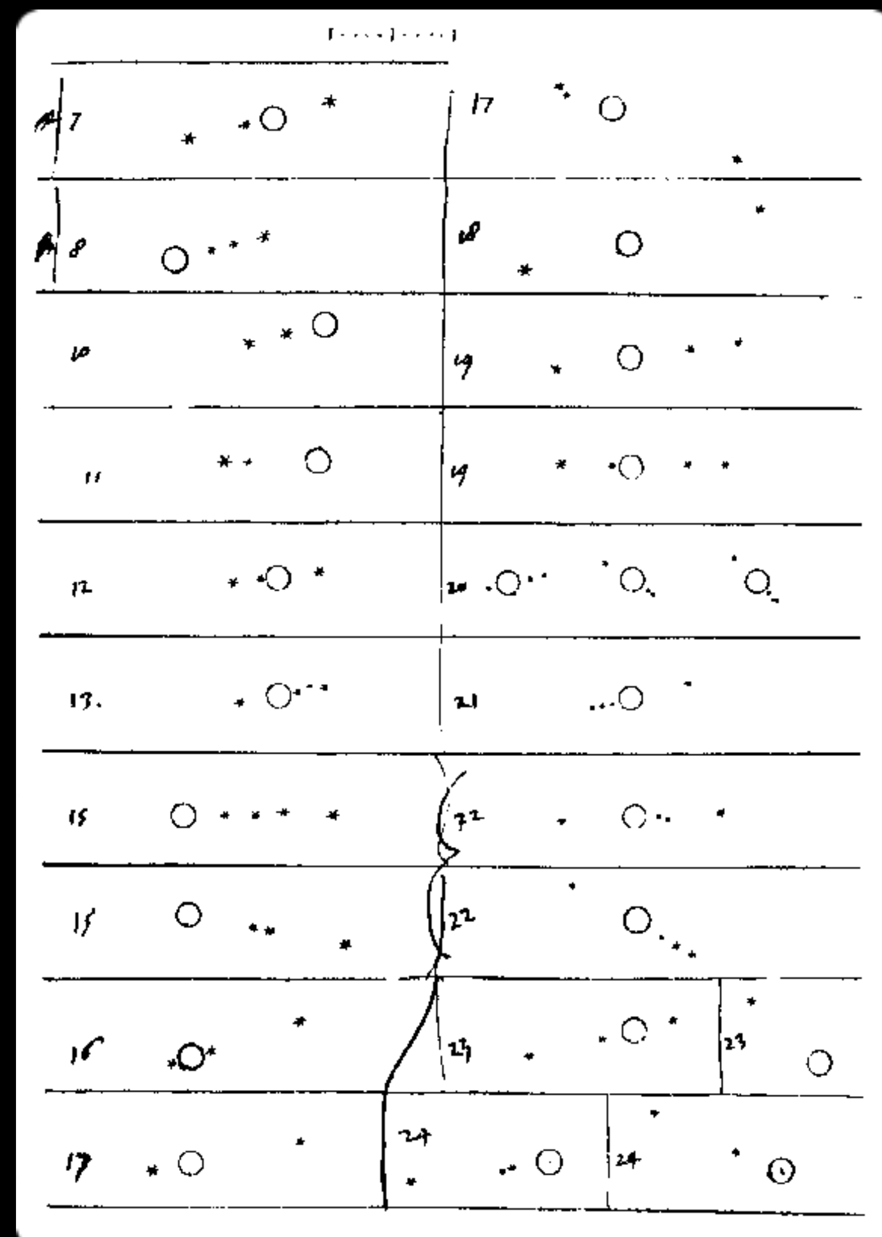


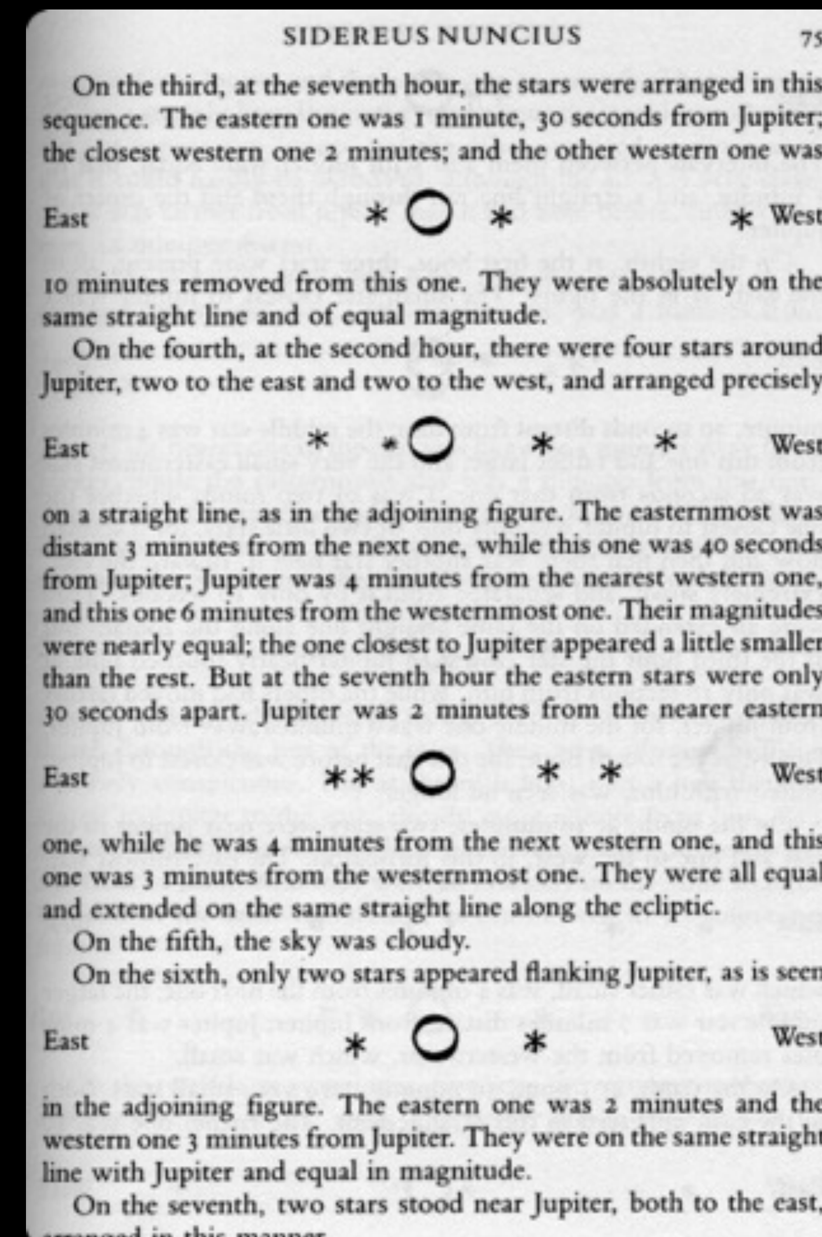
1610:Why Galileo is My Hero

(Explore-Explain-Explore)









SIDEREUS NUNCIUS 75

On the third, at the seventh hour, the stars were arranged in this sequence. The eastern one was 1 minute, 30 seconds from Jupiter; the closest western one 2 minutes; and the other western one was

East * O * * West

10 minutes removed from this one. They were absolutely on the same straight line and of equal magnitude.

On the fourth, at the second hour, there were four stars around Jupiter, two to the east and two to the west, and arranged precisely

East * *O * * West

on a straight line, as in the adjoining figure. The easternmost was distant 3 minutes from the next one, while this one was 40 seconds from Jupiter; Jupiter was 4 minutes from the nearest western one, and this one 6 minutes from the westernmost one. Their magnitudes were nearly equal; the one closest to Jupiter appeared a little smaller than the rest. But at the seventh hour the eastern stars were only 30 seconds apart. Jupiter was 2 minutes from the nearer eastern

East ** O * * West

one, while he was 4 minutes from the next western one, and this one was 3 minutes from the westernmost one. They were all equal and extended on the same straight line along the ecliptic.

On the fifth, the sky was cloudy.

On the sixth, only two stars appeared flanking Jupiter, as is seen

East * O * West

in the adjoining figure. The eastern one was 2 minutes and the western one 3 minutes from Jupiter. They were on the same straight line with Jupiter and equal in magnitude.

On the seventh, two stars stood near Jupiter, both to the east, arranged in this manner.

Notes for & re-productions of Siderius Nuncius

[More later?]

2018: 10QViz.org

≡ MENU

TEN QUESTIONS TO ASK WHEN CREATING A VISUALIZATION

The 10 Questions

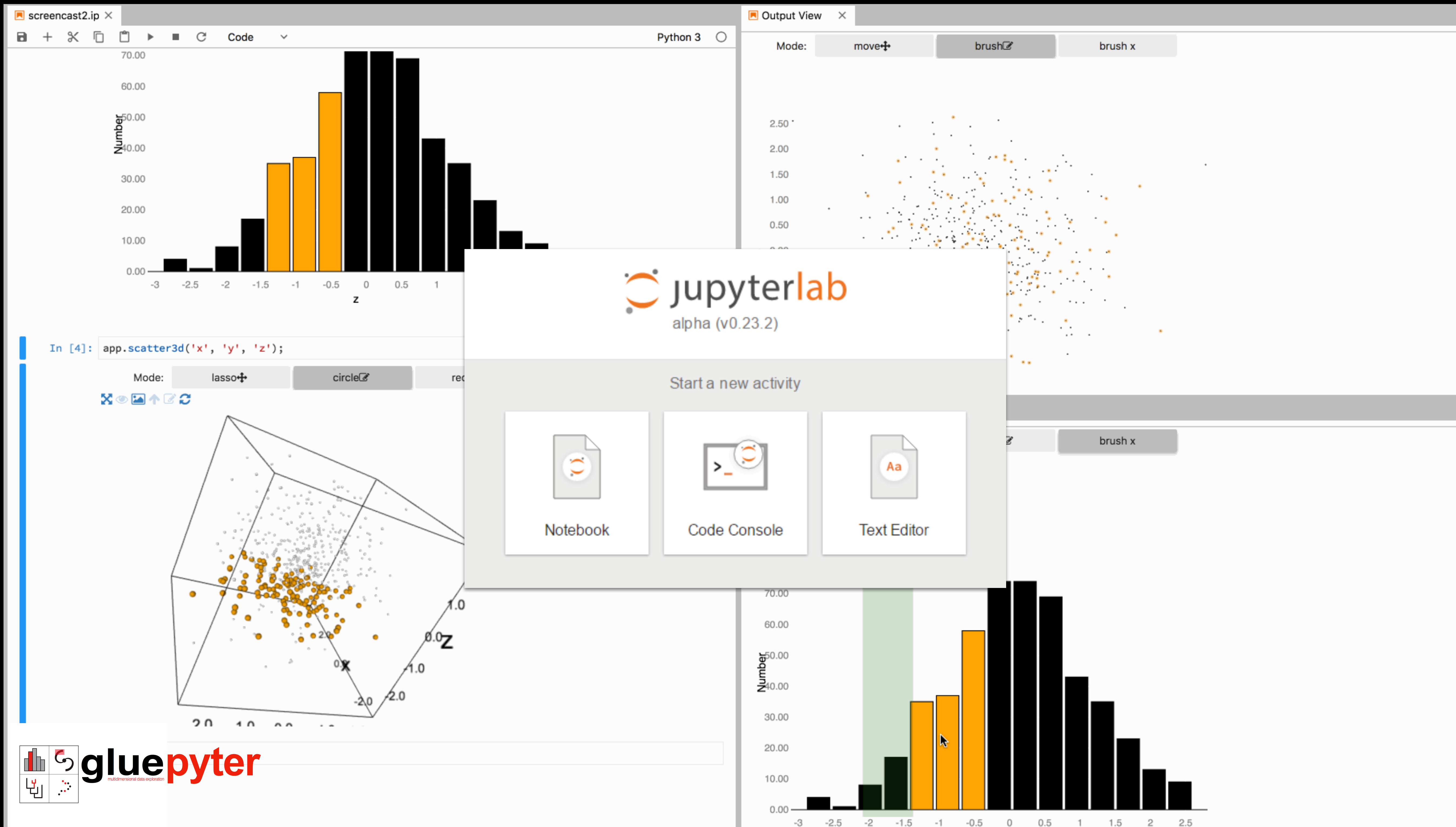
1. **Who** | Who is your audience? How expert will they be about the subject and/or display conventions?
2. **Explore-Explain** | Is your goal to explore, document, or explain your data or ideas, or a combination of these?
3. **Categories** |Do you want to show or explore pre-existing, known, human-interpretable, 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.

2019: glue in the browser



Video courtesy of Maarten Breddels, consulting developer

What anyone can do right now

What computer scientists do

What scientists do

What we should do?

3D PDF 2009

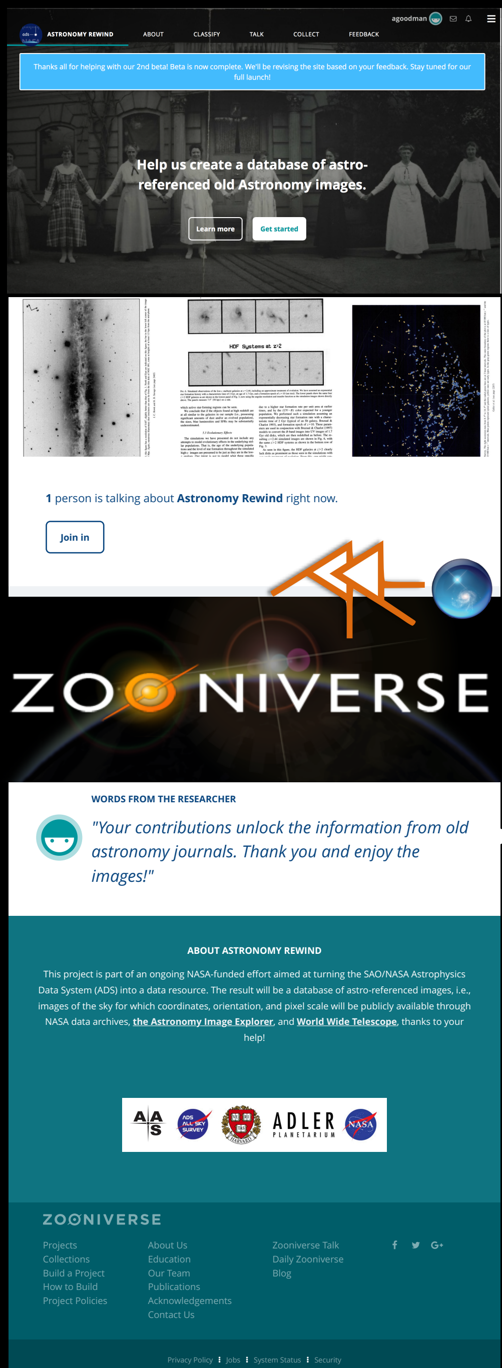
PotF 2015

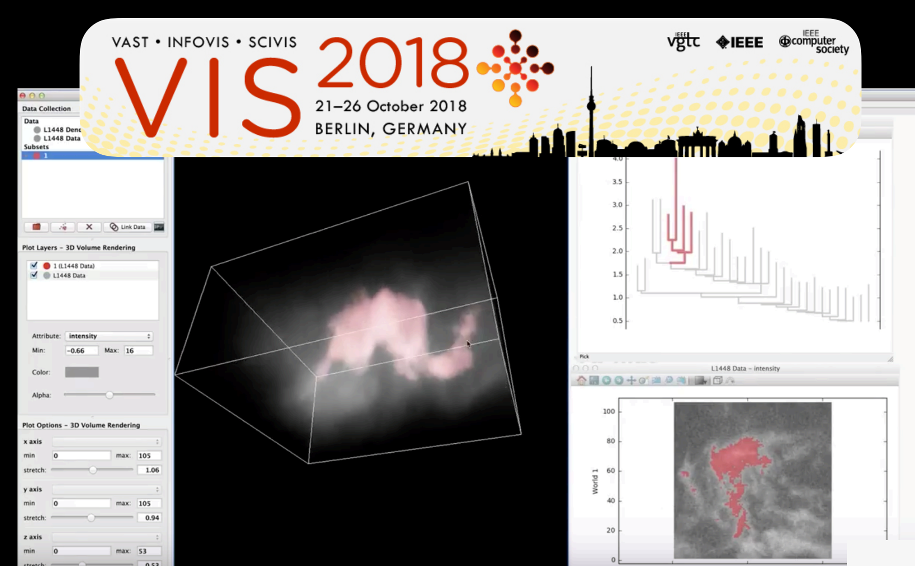
10QViz 2018

glupyter 2019

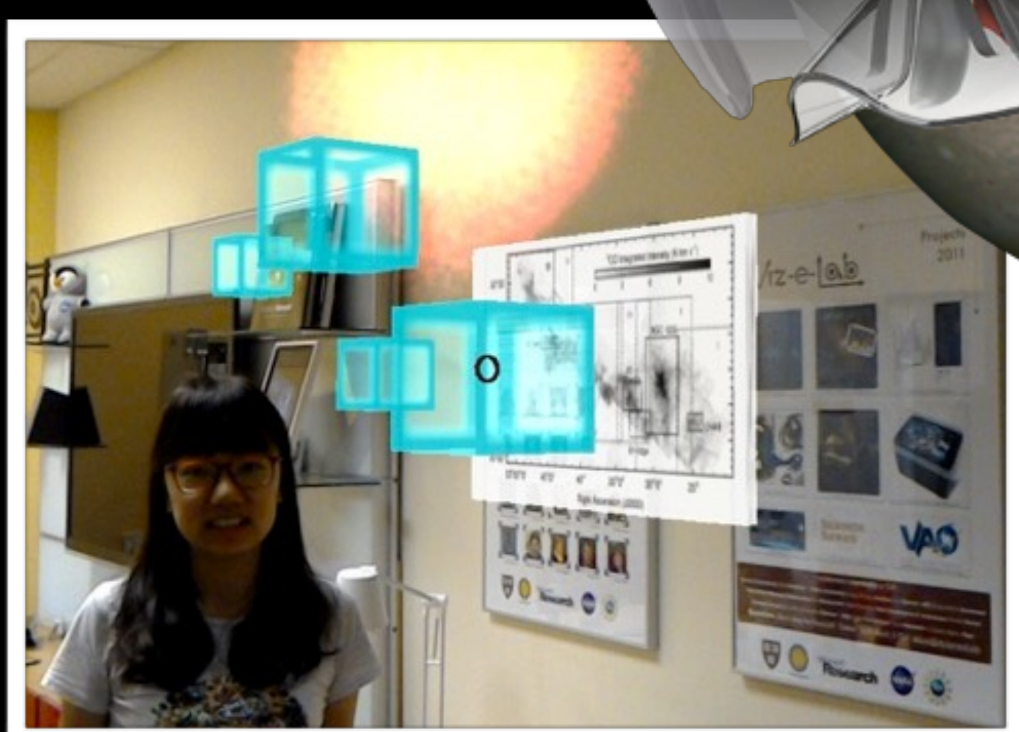
+ more, together

+ more, together



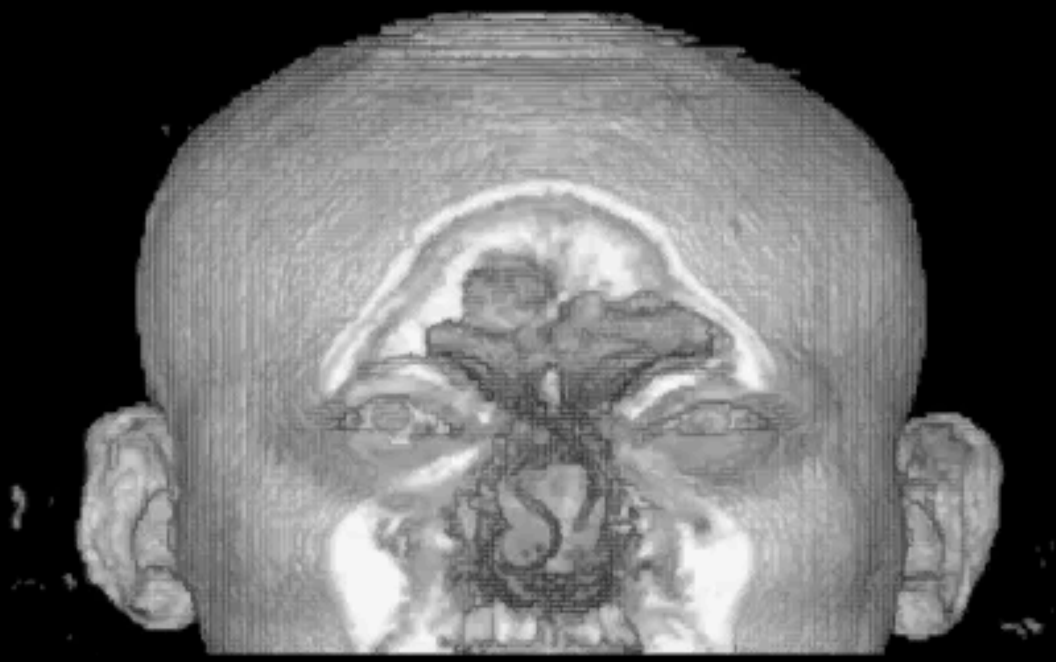


The challenge of 3D Selection



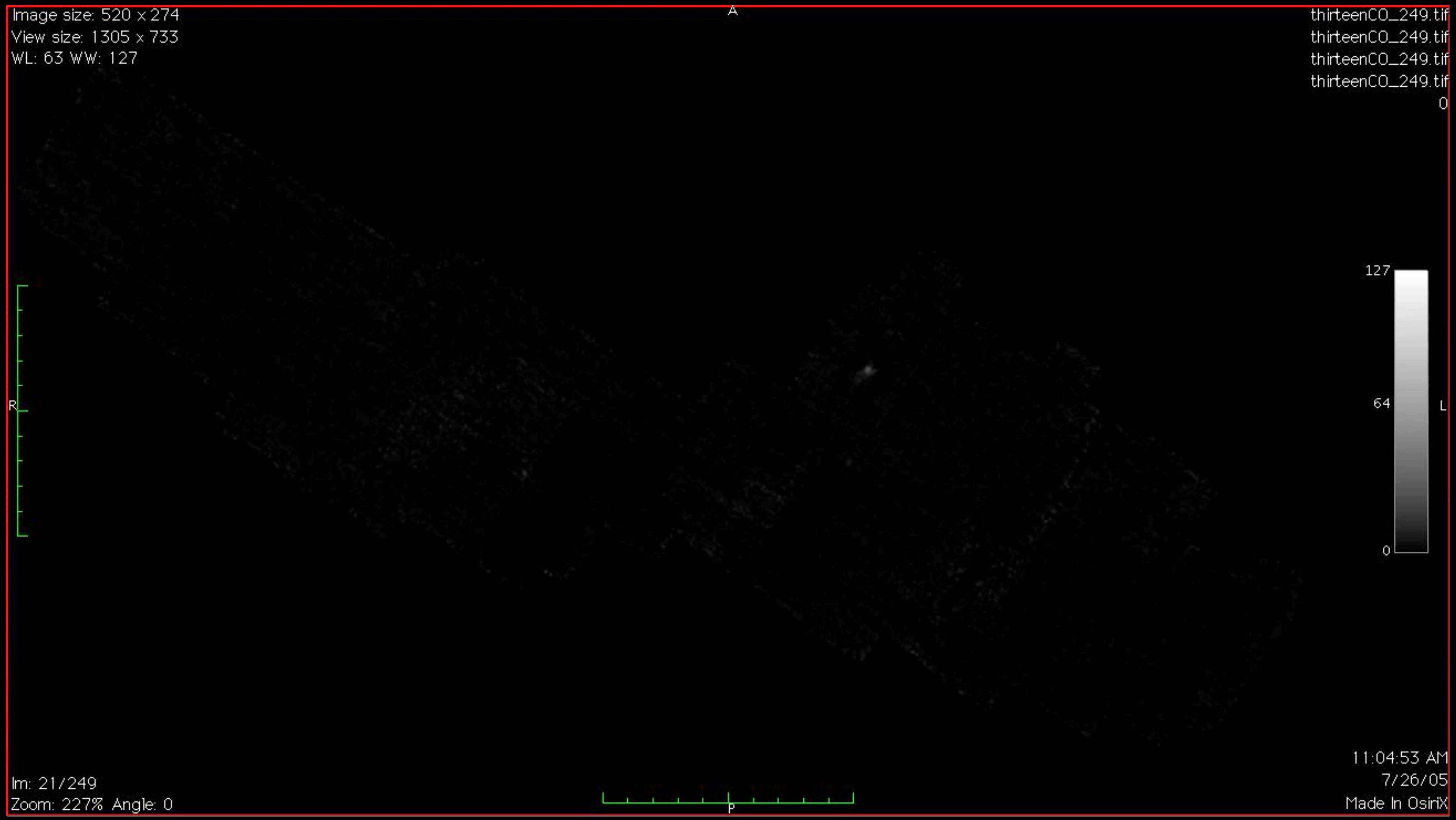
ASTRONOMICAL MEDICINE

"KEITH"



"z" is depth into head

"PERSEUS"



"z" is line-of-sight velocity

ASTRONOMICAL MEDICINE

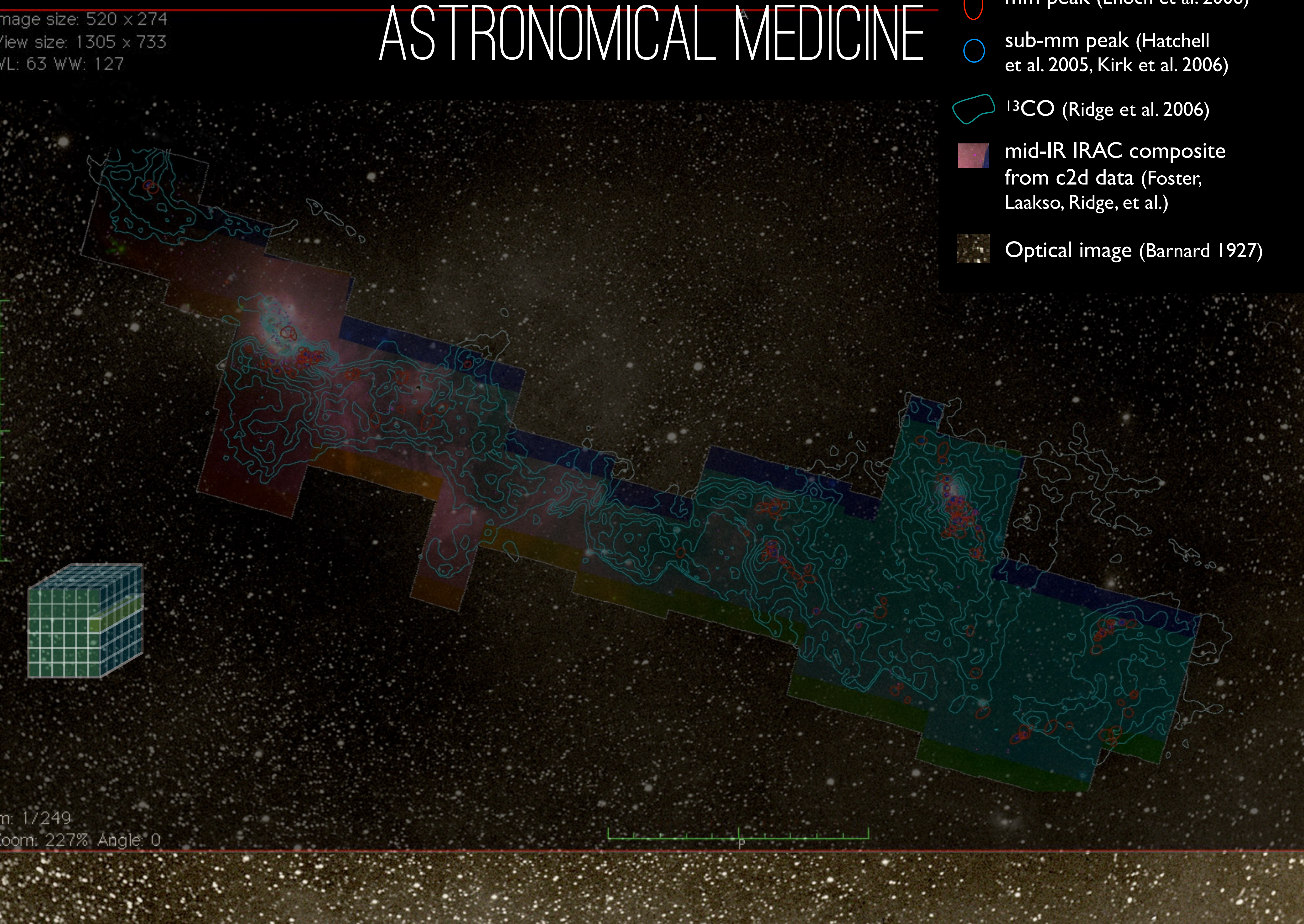
mm peak (Enoch et al. 2006)

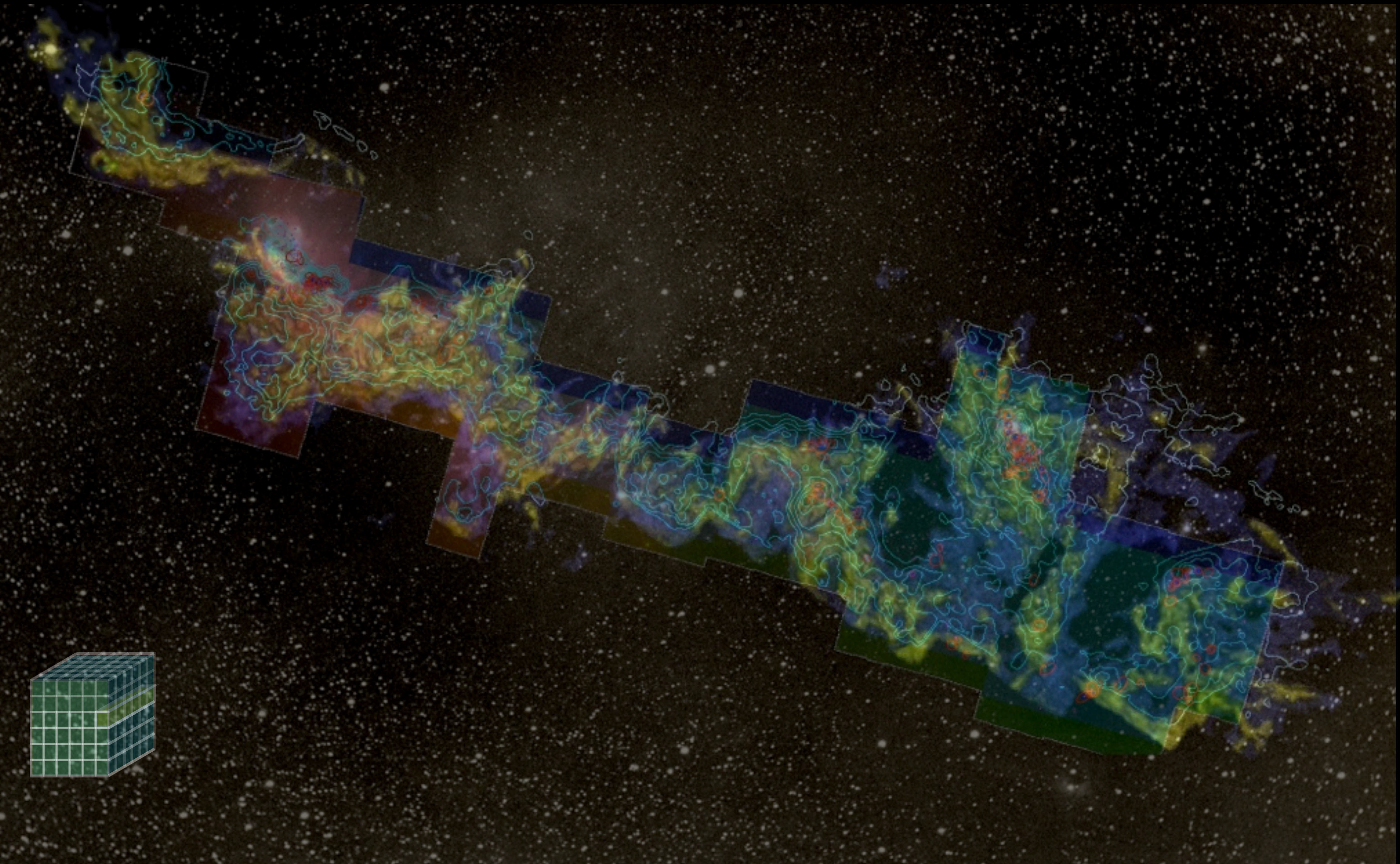
sub-mm peak (Hatchell et al. 2005, Kirk et al. 2006)

^{13}CO (Ridge et al. 2006)

mid-IR IRAC composite from c2d data (Foster, Laakso, Ridge, et al.)

Optical image (Barnard 1927)

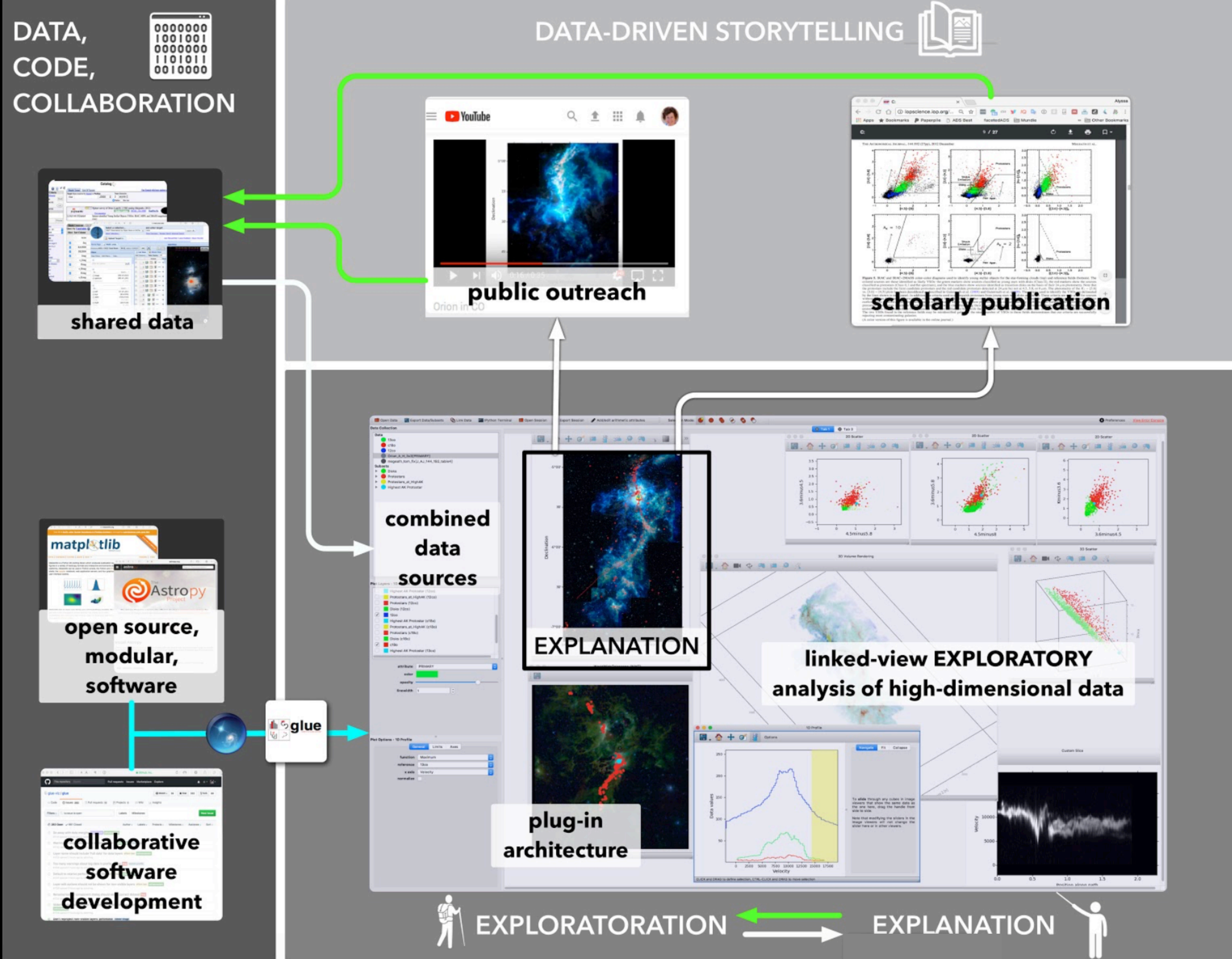




3D Viz made with VolView

COMPLETE

extra slides


DATA,
CODE,
COLLABORATION
DATA-DRIVEN STORYTELLING
shared data
public outreach
scholarly publication
combined
data
sources
EXPLANATION
linked-view EXPLORATORY
analysis of high-dimensional data
plug-in
architecture
open source,
modular,
software
collaborative
software
development
glue
matplotlib
Astropy
EXPLORATORATION
EXPLANATION

glueviz.org github.com/glue-viz glueviz.slack.com

glue

multidimensional data exploration

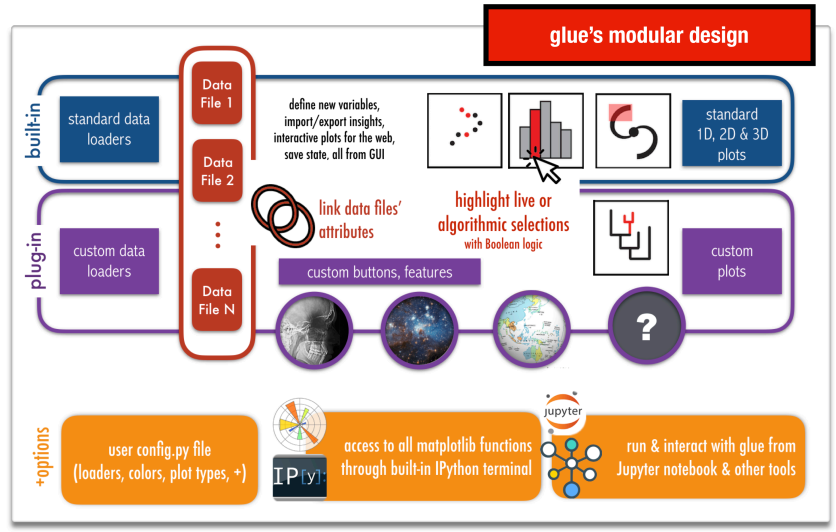
Linked-View Exploratory Visualization of High-Dimensional Data, for Everyone

Alyssa **Goodman** (PI, Harvard)
Michellle **Borkin** (PI, Northeastern)
Thomas **Robitaille** (Lead Architect)

The **glue** project was founded in 2012, with funding from **NASA**'s **James Webb Space Telescope** (JWST) project. NASA contracts continue to support development of JWST-related (Astronomy) functionality.

Beginning in 2017, glue has also been funded by the **National Science Foundation**, under SI2-SSE 1739657/1740229: *Collaborative Research: A sustainable future for the glue multi-dimensional linked data visualization package.* The goal of the NSF SSE funding is to expand glue's functionality into domains beyond its traditional strengths in Astronomy and Medicine, by broadening both its user and developer communities.

All glue code is Open Source, at github.com/glue-viz



Linked Visualizations

With Glue, users can create scatter plots, histograms and images (2D and 3D) of their data. Glue is focused on the brushing and linking paradigm, where selections in any graph propagate to all others.

Flexible linking across data

Glue uses the logical links that exist between different data sets to overlay visualizations of different data, and to propagate selections across data sets. These links are specified by the user, and are arbitrarily flexible

Full scripting capability

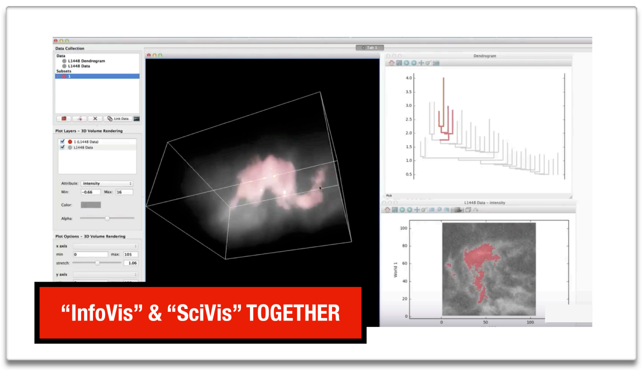
Glue is written in Python, and built on top of its standard scientific libraries (i.e., Numpy, Matplotlib, Scipy). Users can easily integrate their own python code for data input, cleaning, and analysis.

Want to plug-in your project or tool? Consider joining us for glue-con, right after JupyterCon, August 27-29, 2018, at Harvard.

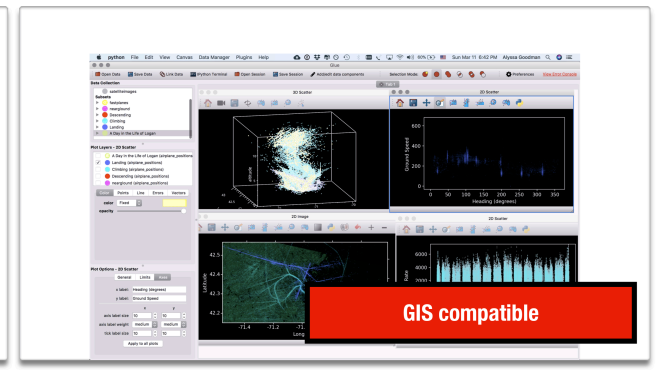
glue-con

2018, Cambridge, MA

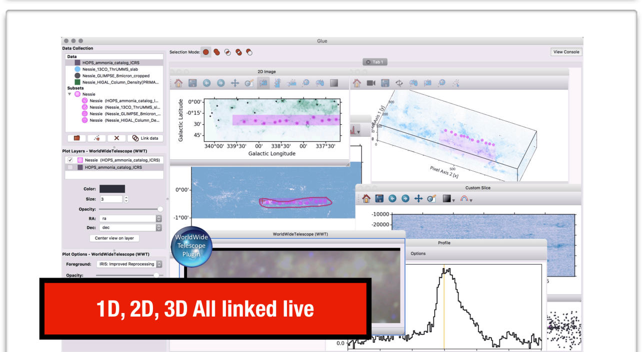
projects.iq.harvard.edu/gluecon



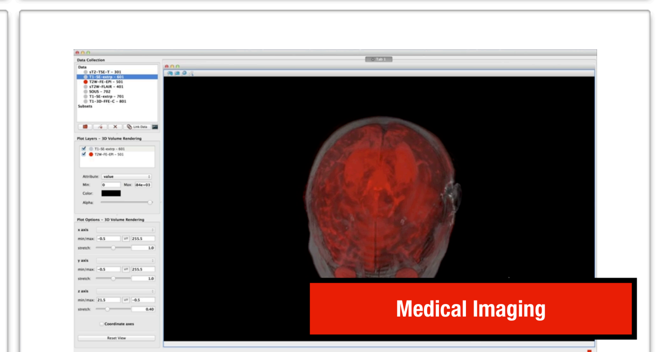
"InfoVis" & "SciVis" TOGETHER



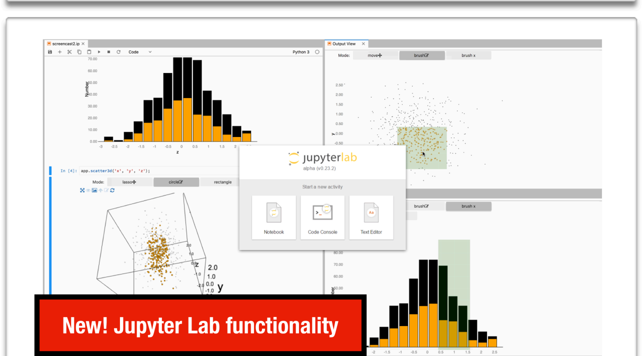
GIS compatible



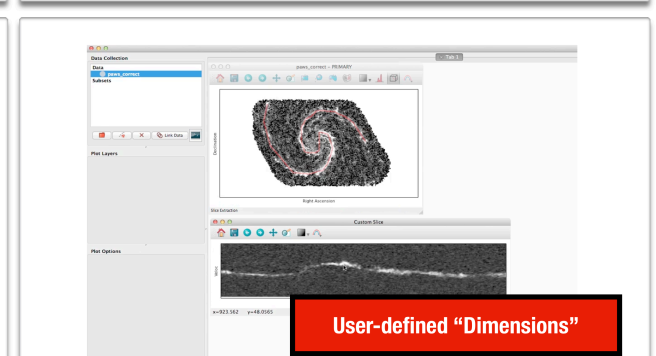
1D, 2D, 3D All linked live



Medical Imaging

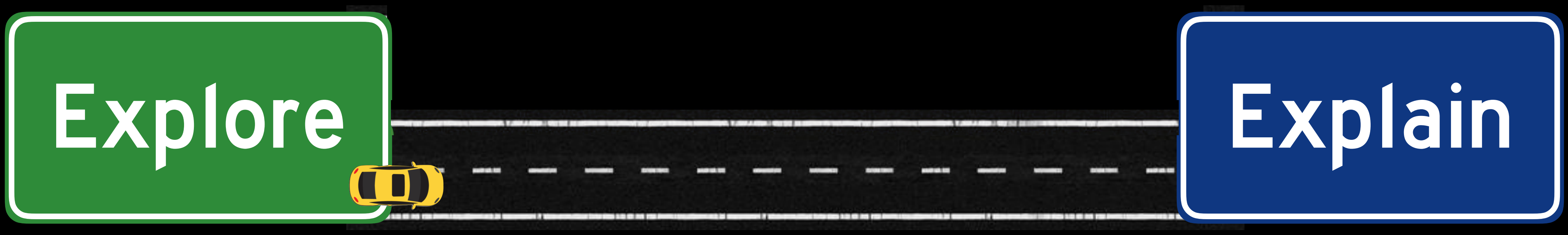


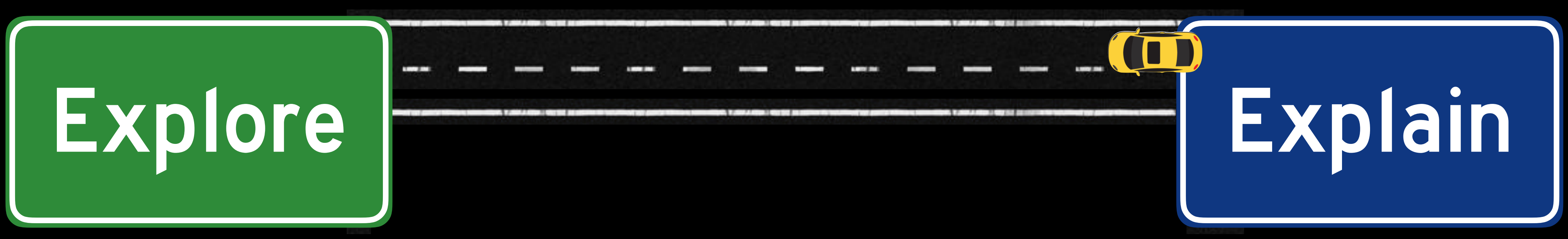
New! Jupyter Lab functionality



User-defined "Dimensions"

Traditionally, travel from exploration to explanation is called "Scholarly Publishing" if its *dry*, and "Public Outreach," if it's *beautiful*.

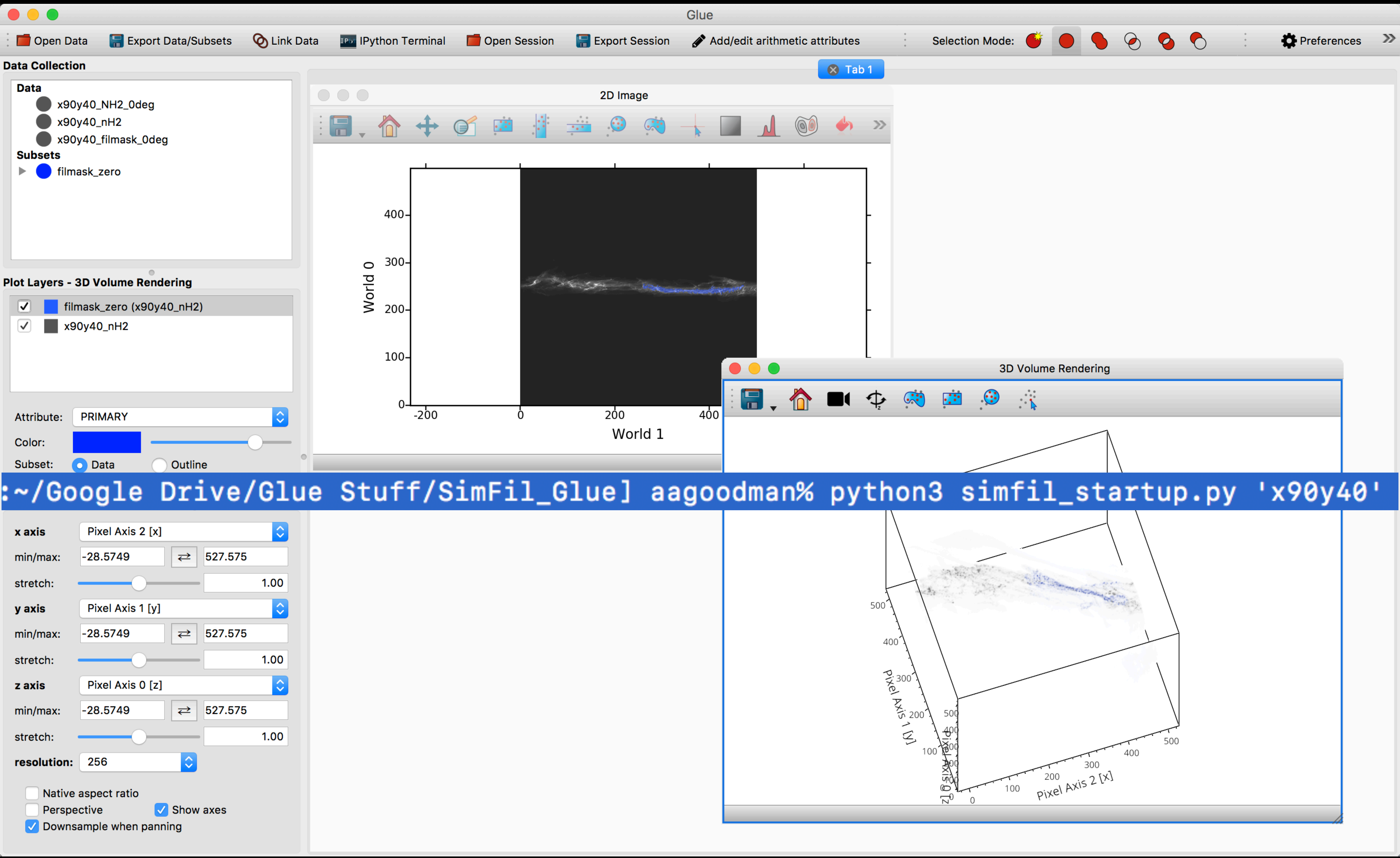




It’s much harder to go the other way.


Explore
And, the *best* roads are two-way.
Explain

Scripting




```
:~/Google Drive/Glue Stuff/SimFil_Glue] aagoodman% python3 simfil_startup.py 'x90y40'
```