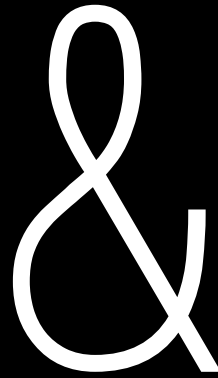


HIGH-DIMENSIONAL  
DATA  
VISUALIZATION



THE "PAPER"  
OF THE  
FUTURE

**Alyssa A. Goodman • Harvard-Smithsonian Center for Astrophysics**

# COLLABORATORS



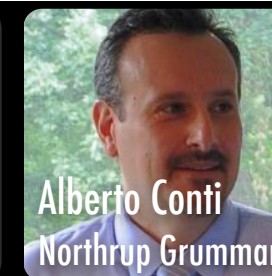
Chris Beaumont,  
CfA



Thomas Boch,  
CDS



Michelle Borkin  
Harvard



Alberto Conti  
Northrup Grumman



Yuan-Sen Ting,  
CfA



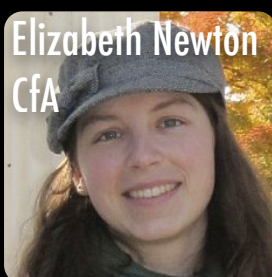
Jonathan Fay  
Microsoft Research



David Hogg  
NYU



Chris Lintott  
Zooniverse, Oxford



Elizabeth Newton  
CfA



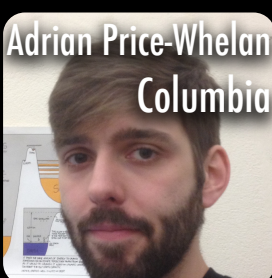
August Munech  
CfA



Josh Peek  
Columbia



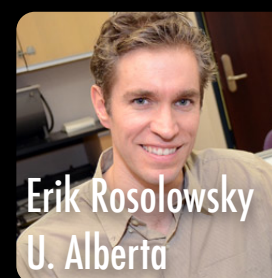
Alberto Pepe  
CfA, Authorea



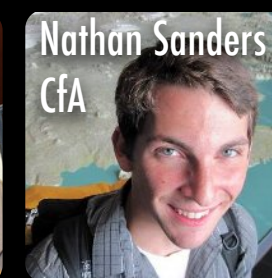
Adrian Price-Whelan  
Columbia



Thomas Robitaille  
MPA



Erik Rosolowsky  
U. Alberta



Nathan Sanders  
CfA



Rob Simpson  
Zooniverse, Oxford



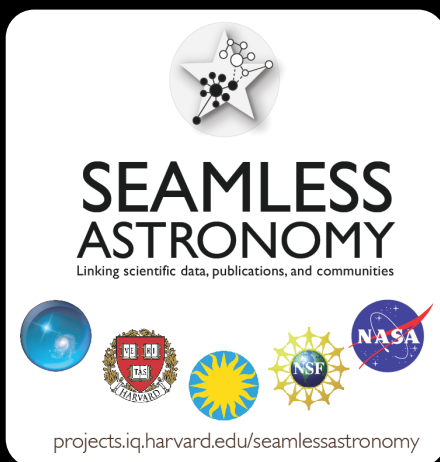
Matt Turk  
Columbia



Pat Udomprasert  
CfA

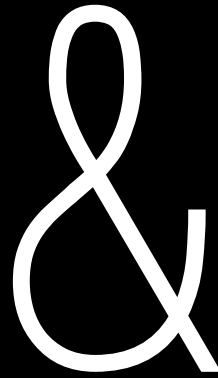


Curtis Wong  
Microsoft Research



...including ADS team (Alberto Accomazzi, Michael Kurtz, Edwin Henneken, et al.) and Wolbach Library staff (Christopher Erdmann et al.)

HIGH-DIMENSIONAL  
DATA  
VISUALIZATION



THE "PAPER"  
OF THE  
FUTURE

**Alyssa A. Goodman • Harvard-Smithsonian Center for Astrophysics**

# RELATIVE STRENGTHS

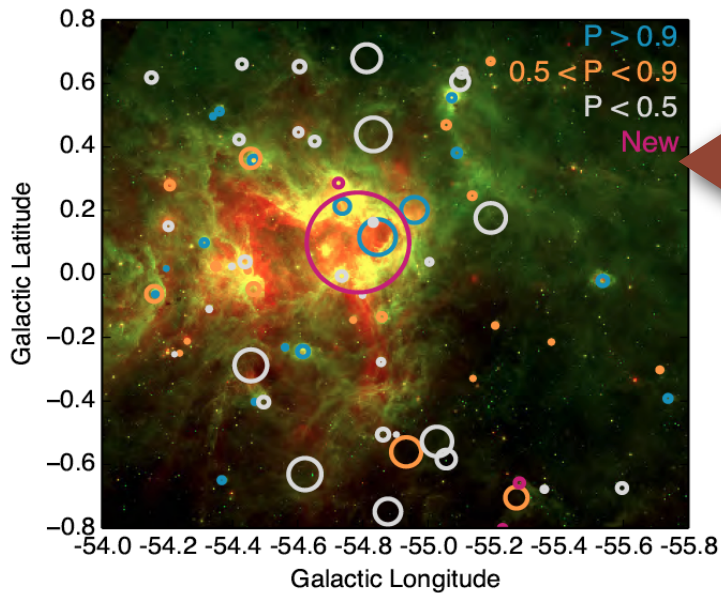
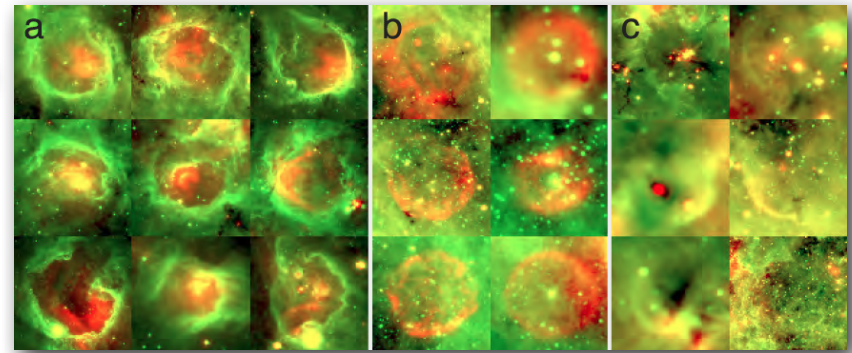


**Pattern Recognition**  
**Creativity**




**Calculations**

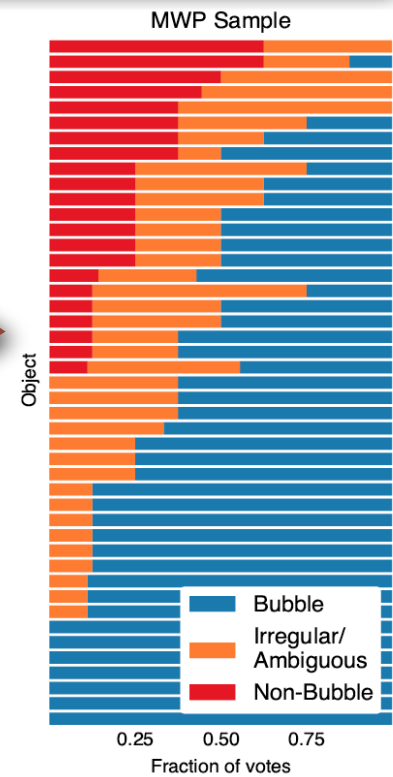
# HUMAN-AIDED COMPUTING



machine-learning algorithm (Brut)



A central box labeled "machine-learning algorithm (Brut)" with a laptop icon below it, indicating the computational processing stage.



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

# VISUALIZATION RESEARCH TODAY



posted by user: [mark\\_s](#) | 471 views | tracked by 6 users: [\[display\]](#) [Add to My List](#) [I'm Organizer](#) [Modify](#)

VIS 2015 : IEEE Visualization: Visual Analytics Science and Technology (VAST), Information Visualization (InfoVis) & Scientific Visualization (SciVis)

[f](#) [t](#) [in](#) [s](#)

Link: <http://www.ieeevis.org/>

When	Oct 25, 2015 - Oct 30, 2015
Where	Chicago, USA
Submission Deadline	Mar 31, 2015

Call For Papers

TBA

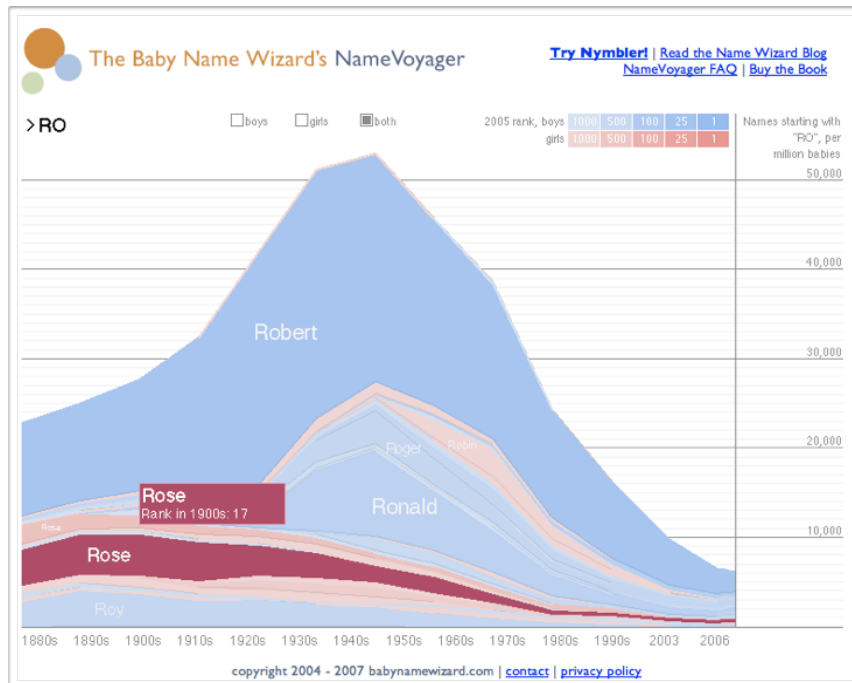
(information from [http://infovis-wiki.net/index.php?title=VIS\\_2015](http://infovis-wiki.net/index.php?title=VIS_2015))

*"My first reading of your proposal suggests that it is more directed at scientific visualization, computer vision, signal processing and the generation of effective images. My own interests are more on information visualization and visual analytics processes that lead to insights, not images. The description does NOT include mention of multi-variate, time series, tree structures, or networks, which form the bulk of my work. Also, I assume you do not wish to cover infographics designed for the general public."*

-excerpt from an email about attending a workshop on "the intersection between human and algorithmic approaches to finding patterns in visual data, within the context of real-world problems" from one prominent visualization researcher to another, 2012

# INFOVIS

**Information visualization** is the study of (interactive) visual representations of abstract data to reinforce human cognition. The abstract data include both numerical and non-numerical data, such as text and geographic information. However, information visualization differs from scientific visualization: "it's infovis [information visualization] when the spatial representation is chosen, and it's scivis [scientific visualization] when the spatial representation is given". (Wikipedia, 2014)

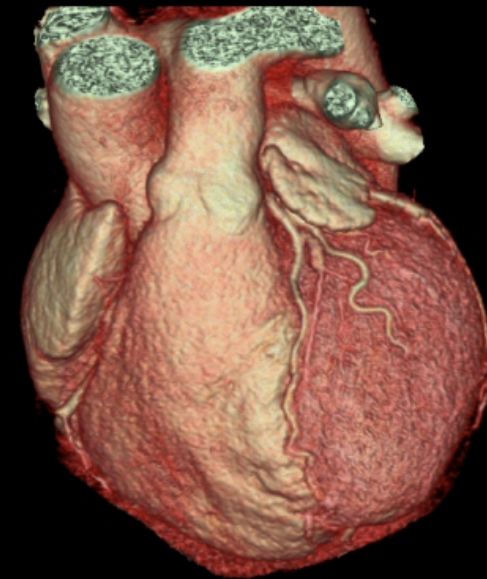


<http://www.babynamewizard.com/namevoyager/Inv0105.html>

created by Martin Wattenberg

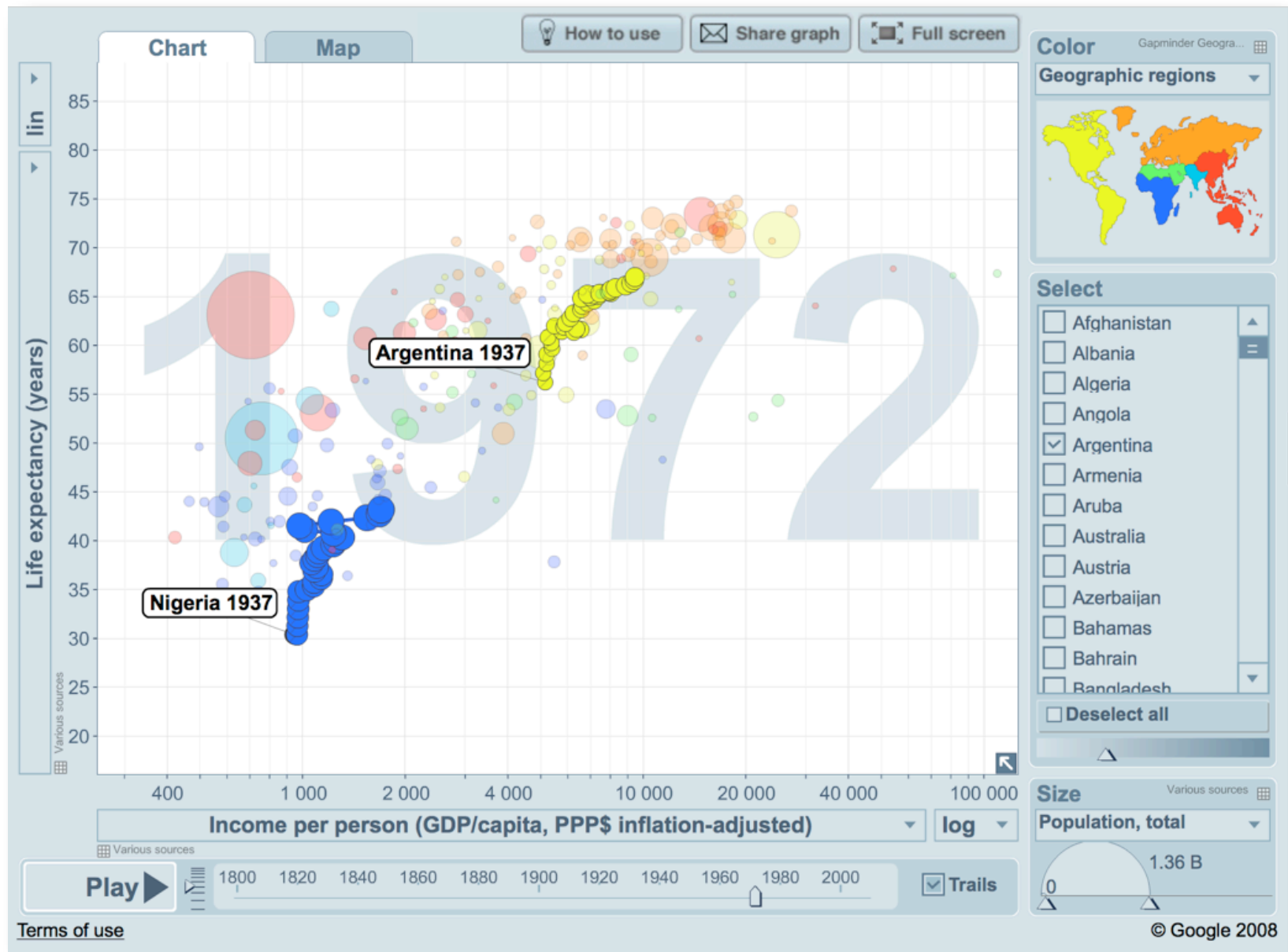
# SCIVIS

**Scientific visualization is an interdisciplinary branch of science.** According to Friendly (2008), it is "primarily concerned with the visualization of three-dimensional phenomena (architectural, meteorological, medical, biological, etc.), where the emphasis is on realistic renderings of volumes, surfaces, illumination sources, and so forth, perhaps with a dynamic (time) component". It is also considered a branch of computer science that is a subset of computer graphics. The purpose of scientific visualization is to graphically illustrate scientific data to enable scientists to understand, illustrate, and glean insight from their data. (Wikipedia 2014)



[http://www.uphs.upenn.edu/news/News\\_Releases/sep05/64sliceCT.htm](http://www.uphs.upenn.edu/news/News_Releases/sep05/64sliceCT.htm)

# INFOVIS? SCIVIS? VISUAL ANALYTICS?





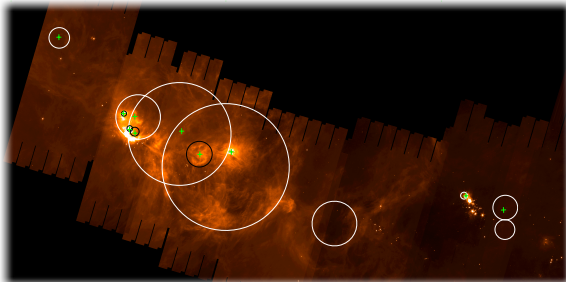
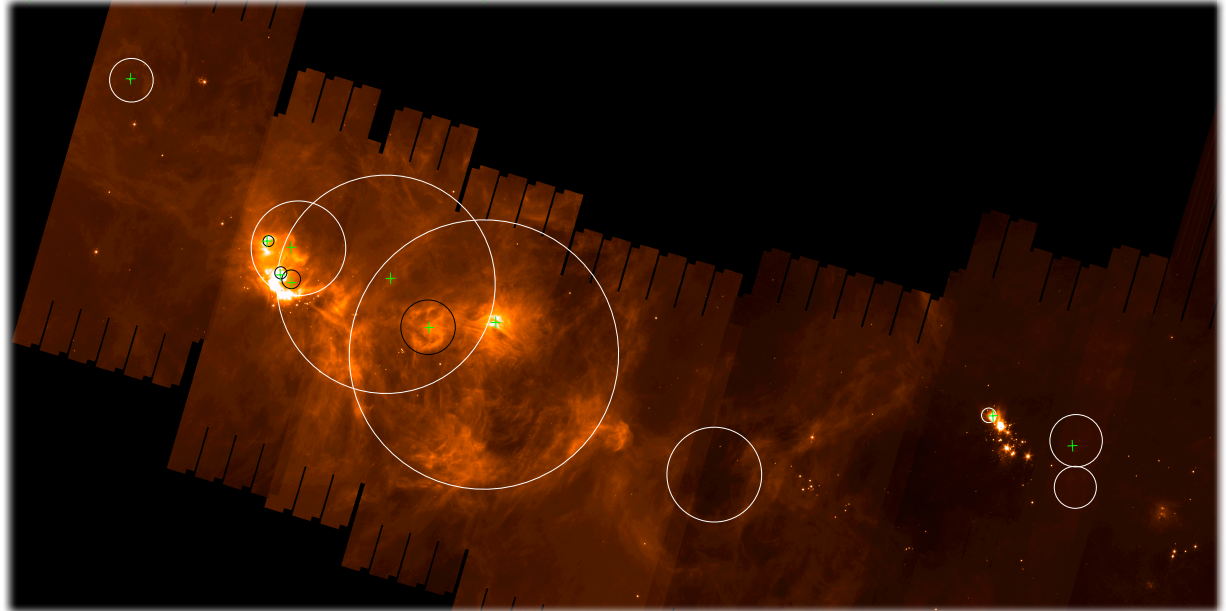
~~INFOVIS? SCIVIS? VISUAL ANALYTICS?~~

**These distinctions need not be made,  
and they harm progress in research.**

DISTINCTIONS  
AMONGST  
VIZ TYPES  
IS UNHELPFUL

e.g. "Shells in Perseus"

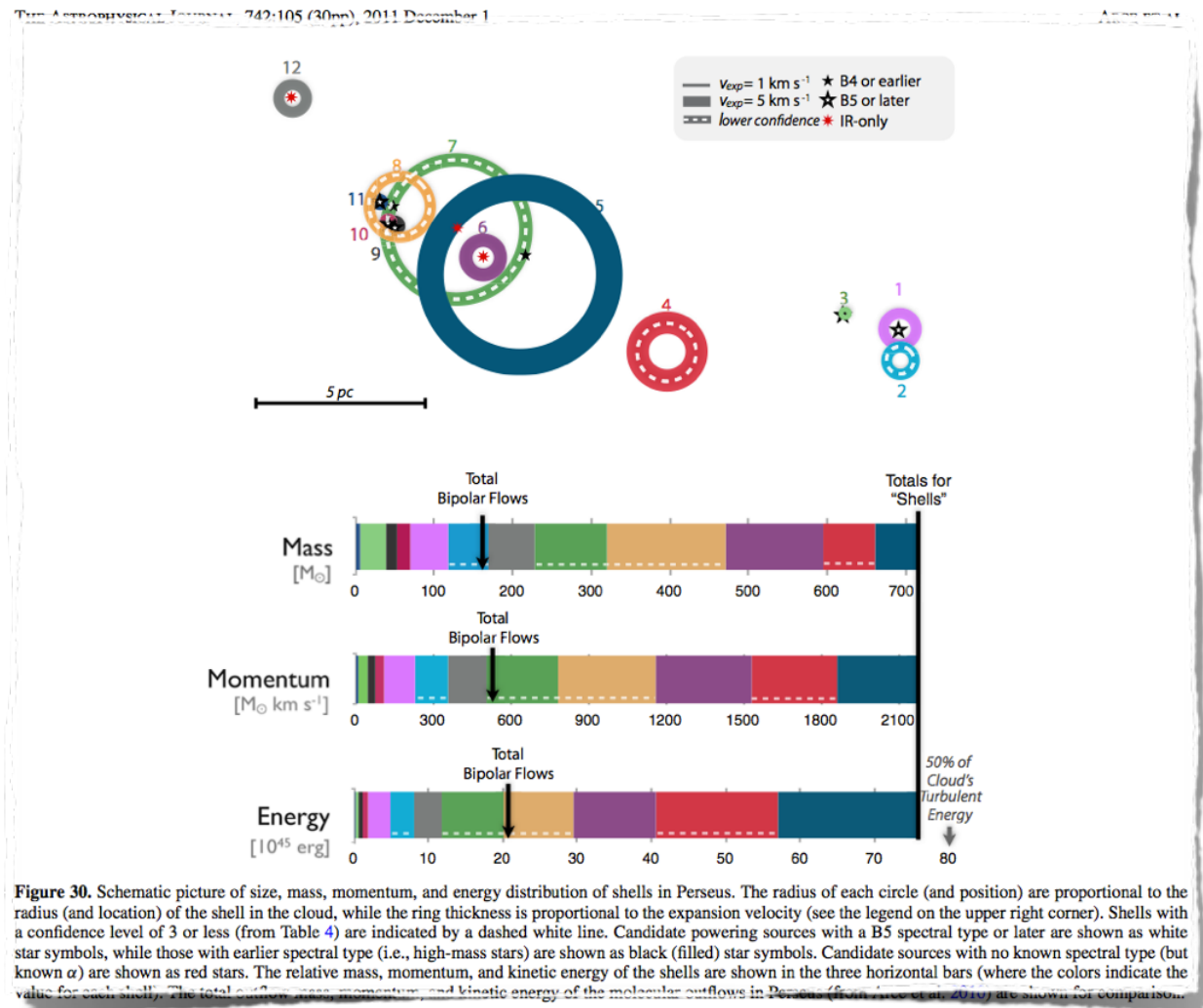
Arce et al. 2011, [10.1088/0004-637X/742/2/105](https://doi.org/10.1088/0004-637X/742/2/105)

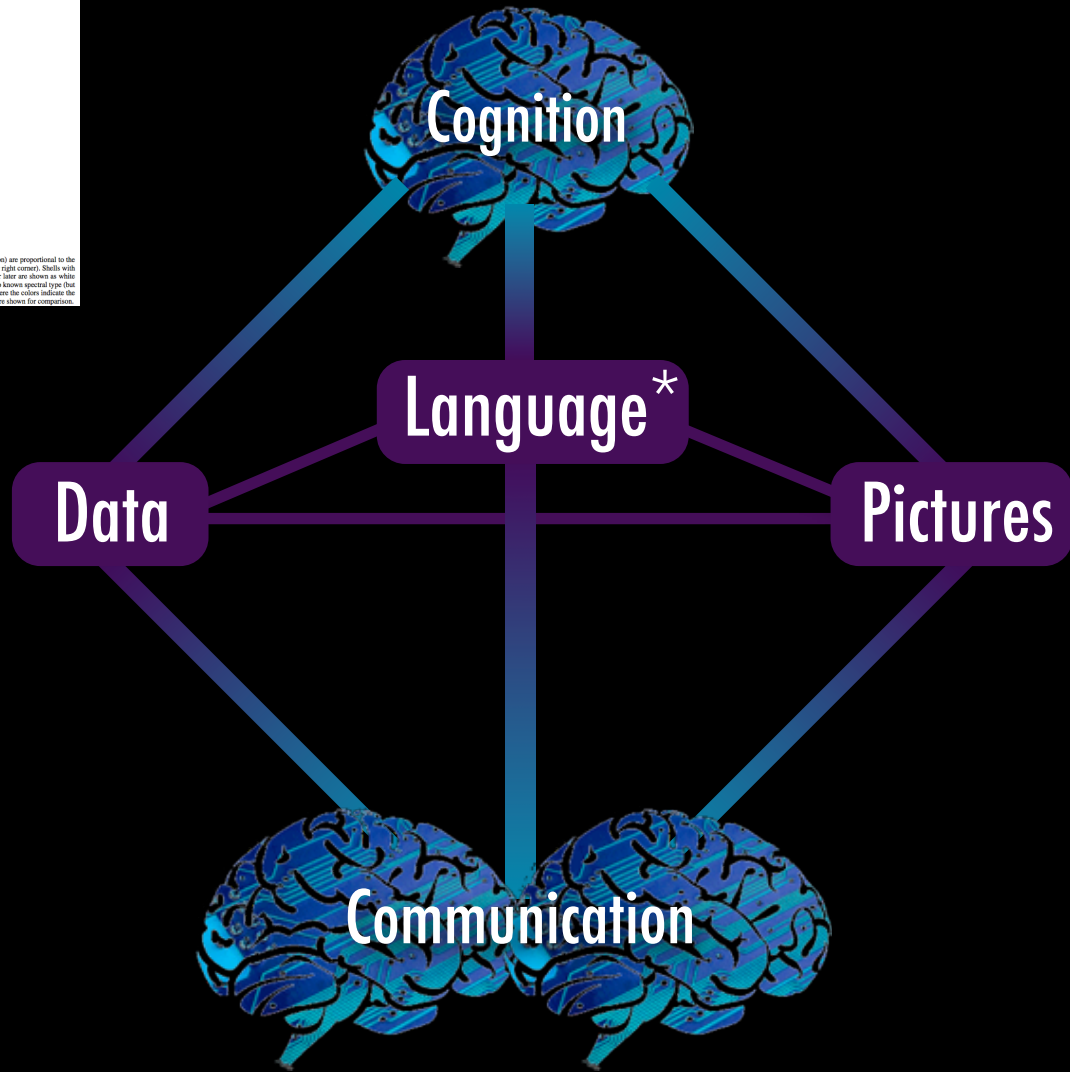
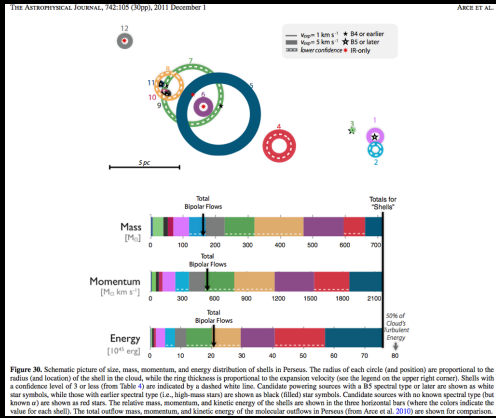


# DISTINCTIONS AMONGST VIZ TYPES IS UNHELPFUL

e.g. "Shells in Perseus"

Arce et al. 2011, [10.1088/0004-637X/742/2/105](https://doi.org/10.1088/0004-637X/742/2/105)





\*"Language" includes words & math

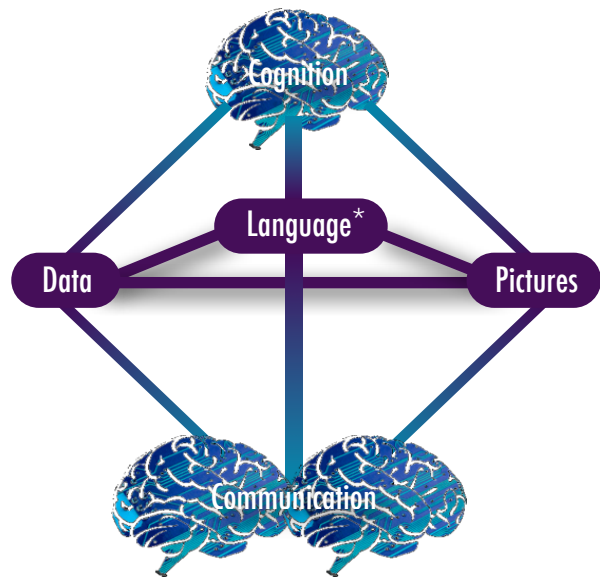
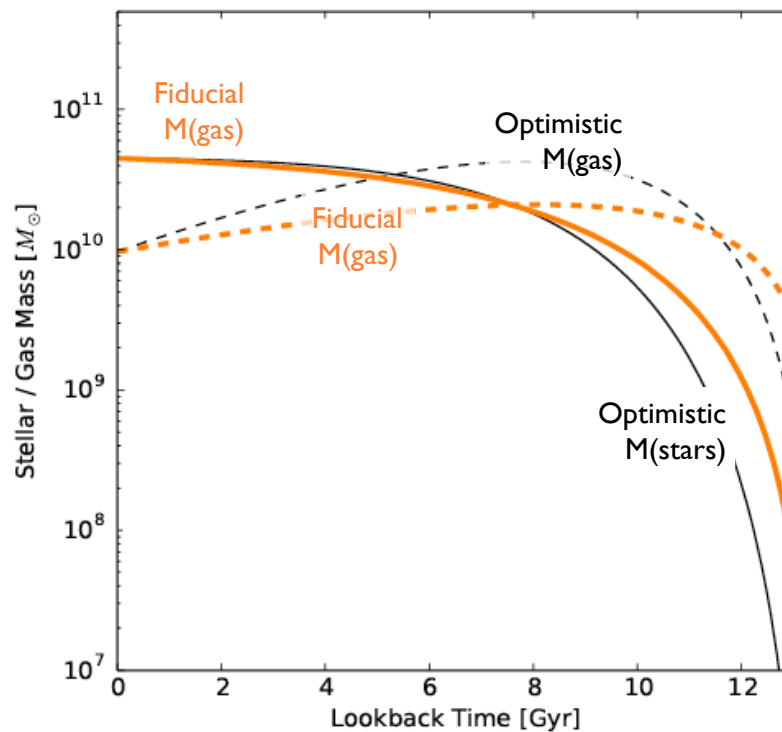
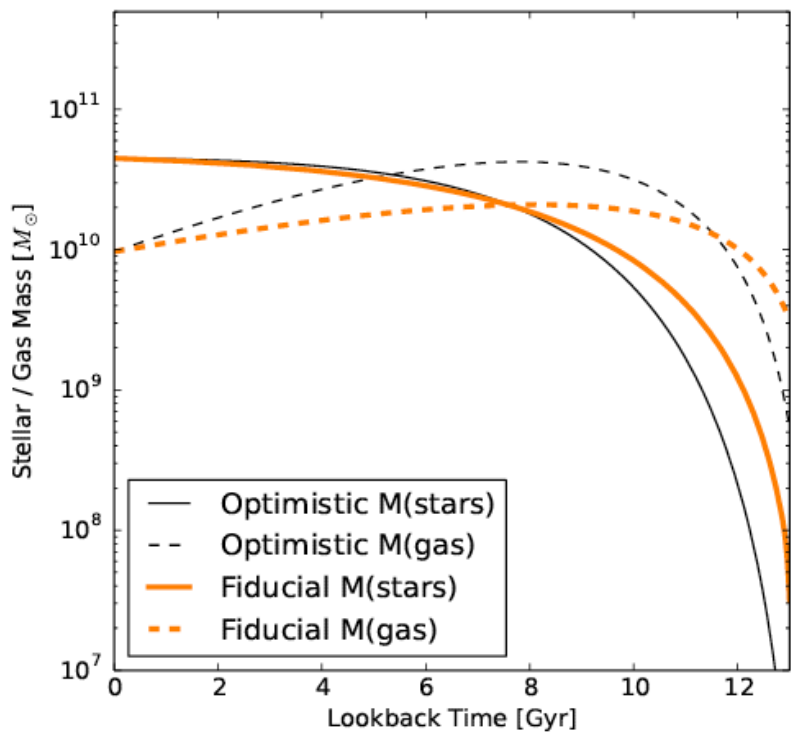
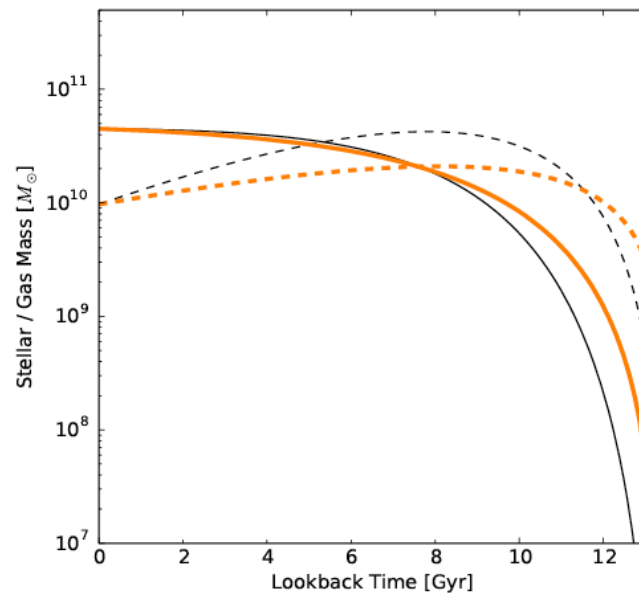


Figure Caption: The solid/solid black line shows the optimistic case for  $M(\text{stars})/M(\text{gas})$ . The orange lines show the same quantities, for the fiducial case.



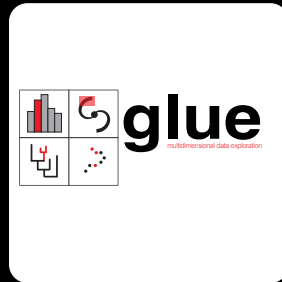


HIGH-DIMENSIONAL  
DATA  
VISUALIZATION

&

THE "PAPER"  
OF THE  
FUTURE





# The Visualization of Astrophysical Data:

Bringing together Science, Art and Education



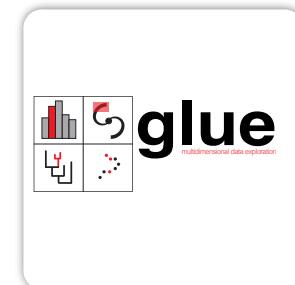
The visualization of real astrophysical data sets is a powerful tool for communicating science to the public and for teaching. The confluence of high quality data sets such as SDSS & WMAP, advances in computational techniques, and the continued march of Moore's law has enabled the use of stunningly beautiful and scientifically accurate images, animations, and interactives in a variety of settings (e.g. TV programs, museums, websites, digital planetaria, magazines, & undergraduate classrooms). Two dimensional, three dimensional, and hyper-dimensional (e.g. color coded 3D data) representations convey large amounts of information in a visceral fashion that can inform both experts and the public. As the data and techniques have progressed the boundaries between art and science have begun to blur and move towards research. This workshop will bring together astrophysicists, visualizers, and educators to discuss the current status and to debate the future direction of astronomical visualization as a tool for research, education, and public outreach.

**Organizers:** Randy Landsberg, Josh Frieman, Andrew Hamilton (CU Boulder), Andrey Kravtsov, Mark SubbaRao, & Alex Szalay (JHU).

2005 IN  
CHICAGO!

thanks Randy, Mark &  
Andrey

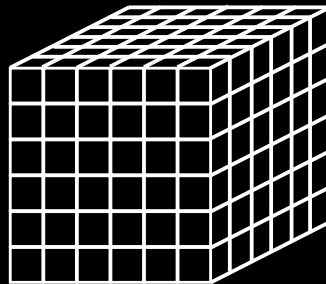
2014 IN  
CHICAGO!  
.astronomy

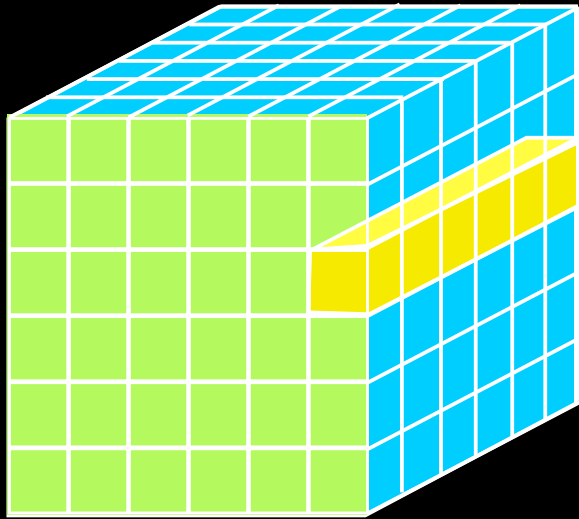




INFOVIS? SCIVIS? VISUAL ANALYTICS?

**These distinctions need not be made,  
and they harm progress in research,  
by ignoring the high-dimensional nature of many data sets.**





"DATA, DIMENSIONS, DISPLAY"

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






**2D:** Faces or Slices = "Images"

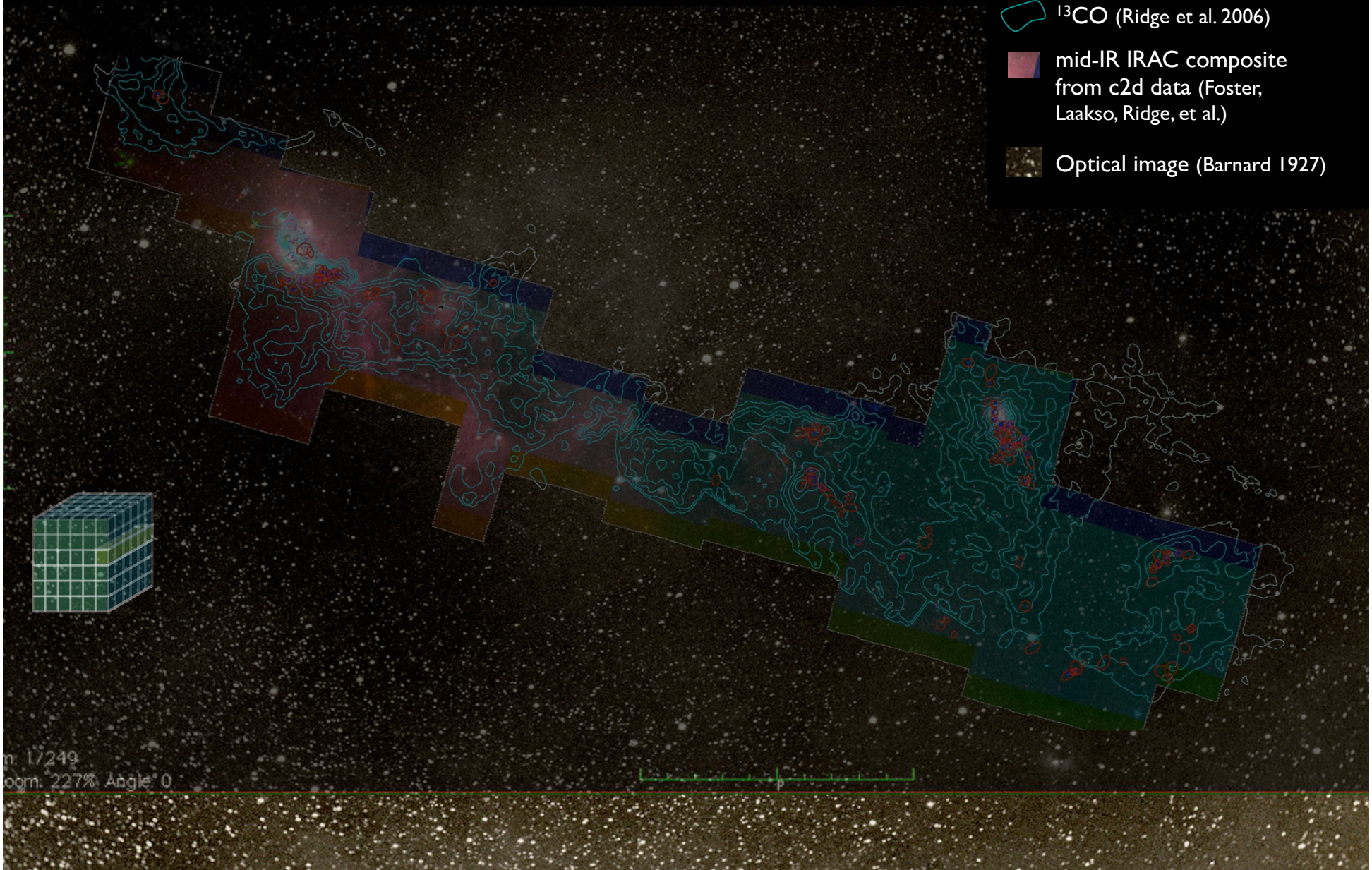
**3D:** Volumes = "3D Renderings", "2D Movies"

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

# DATA, DIMENSIONS, DISPLAY

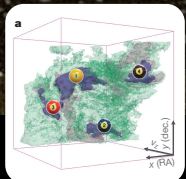
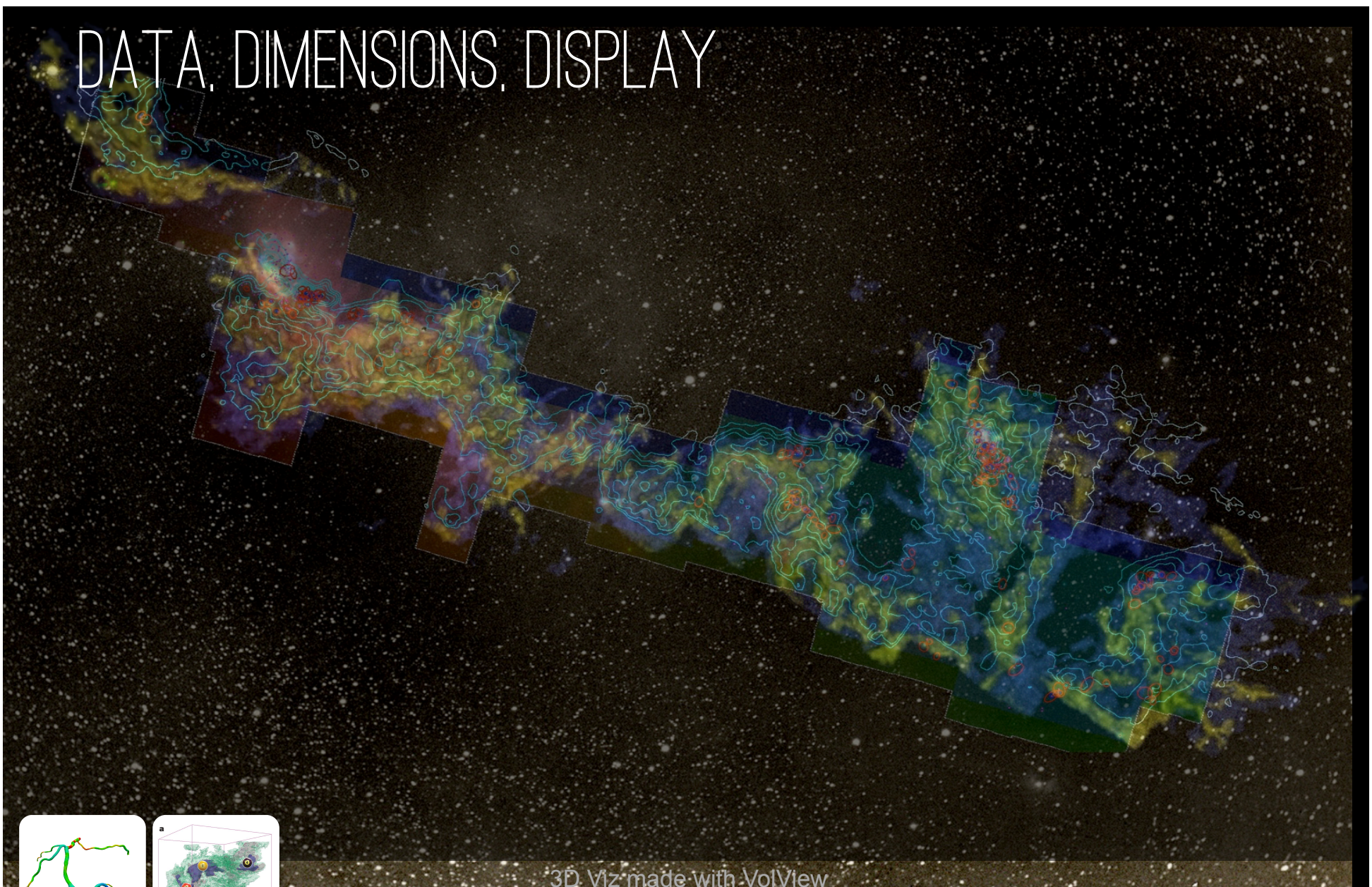
Image size: 510 x 274  
View size: 1305 x 733  
VL: 63 WW: 127

-  mm peak (Enoch et al. 2006)
-  sub-mm peak (Hatchell et al. 2005, Kirk et al. 2006)
-   $^{13}\text{CO}$  (Ridge et al. 2006)
-  mid-IR IRAC composite from c2d data (Foster, Laakso, Ridge, et al.)
-  Optical image (Barnard 1927)



m: 1/249  
Zoom: 227% Angle: 0

# DATA, DIMENSIONS, DISPLAY

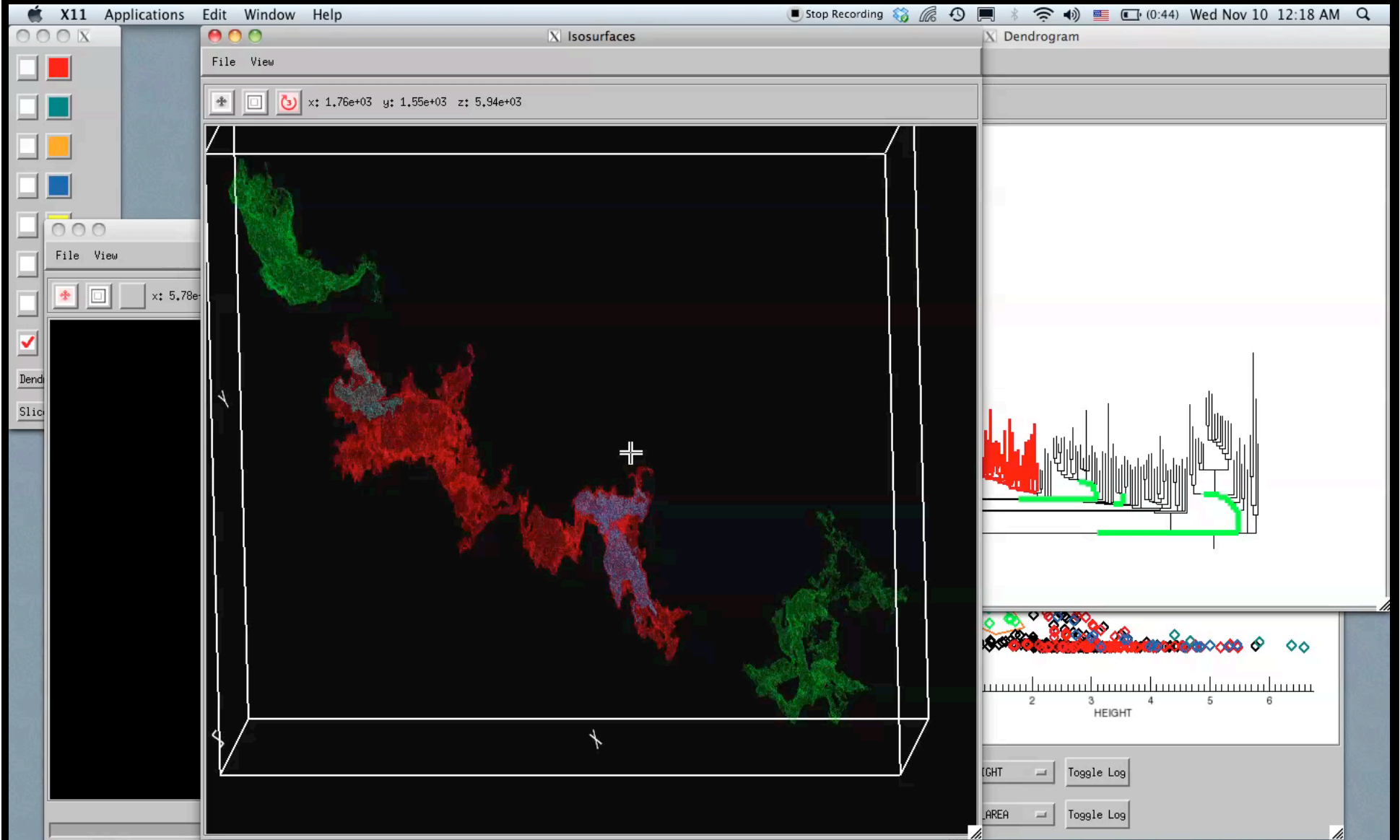


3D Viz made with VolView

Astronomical**Medicine**@**iig**

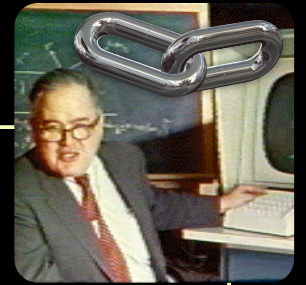
COMPLETE

# INFOVIS? SCIVIS? VISUAL ANALYTICS?

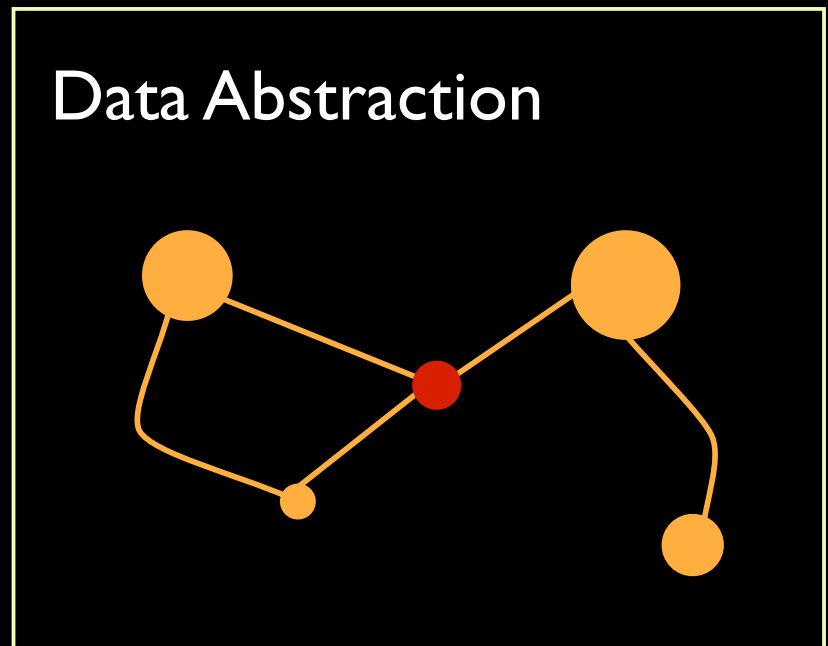
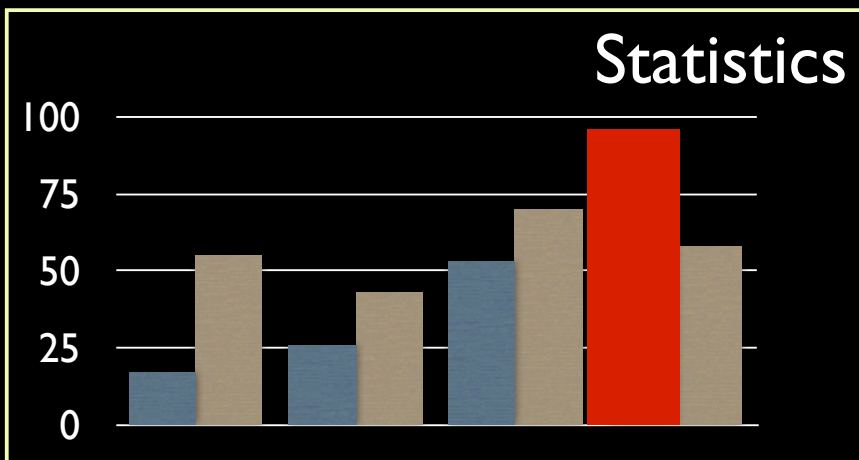
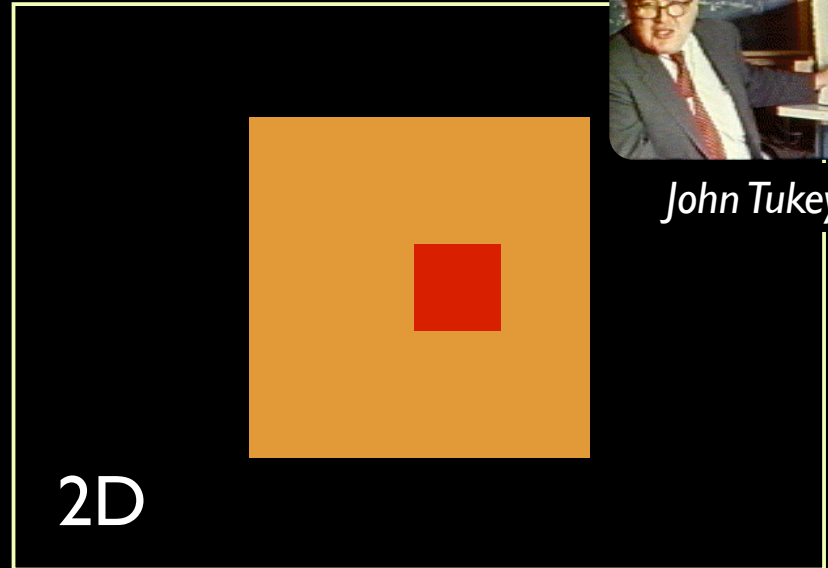
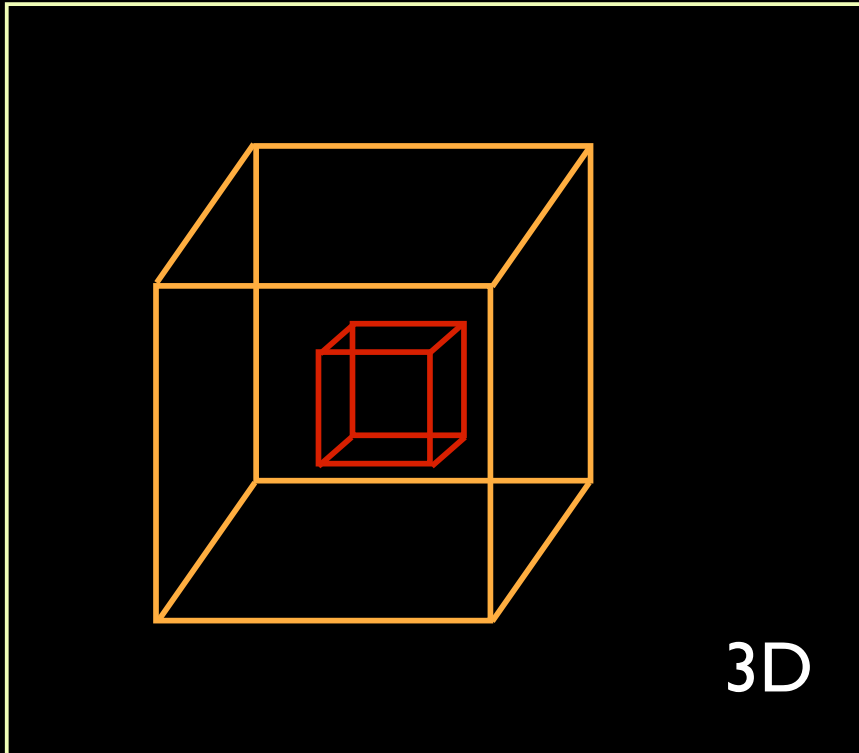


Video & implementation: **Christopher Beaumont**, CfA;  
inspired by AstroMed work of Douglas Alan, Michelle Borkin, AG, Michael Halle, Erik Rosolowsky

# LINKED VIEWS OF HIGH-DIMENSIONAL DATA

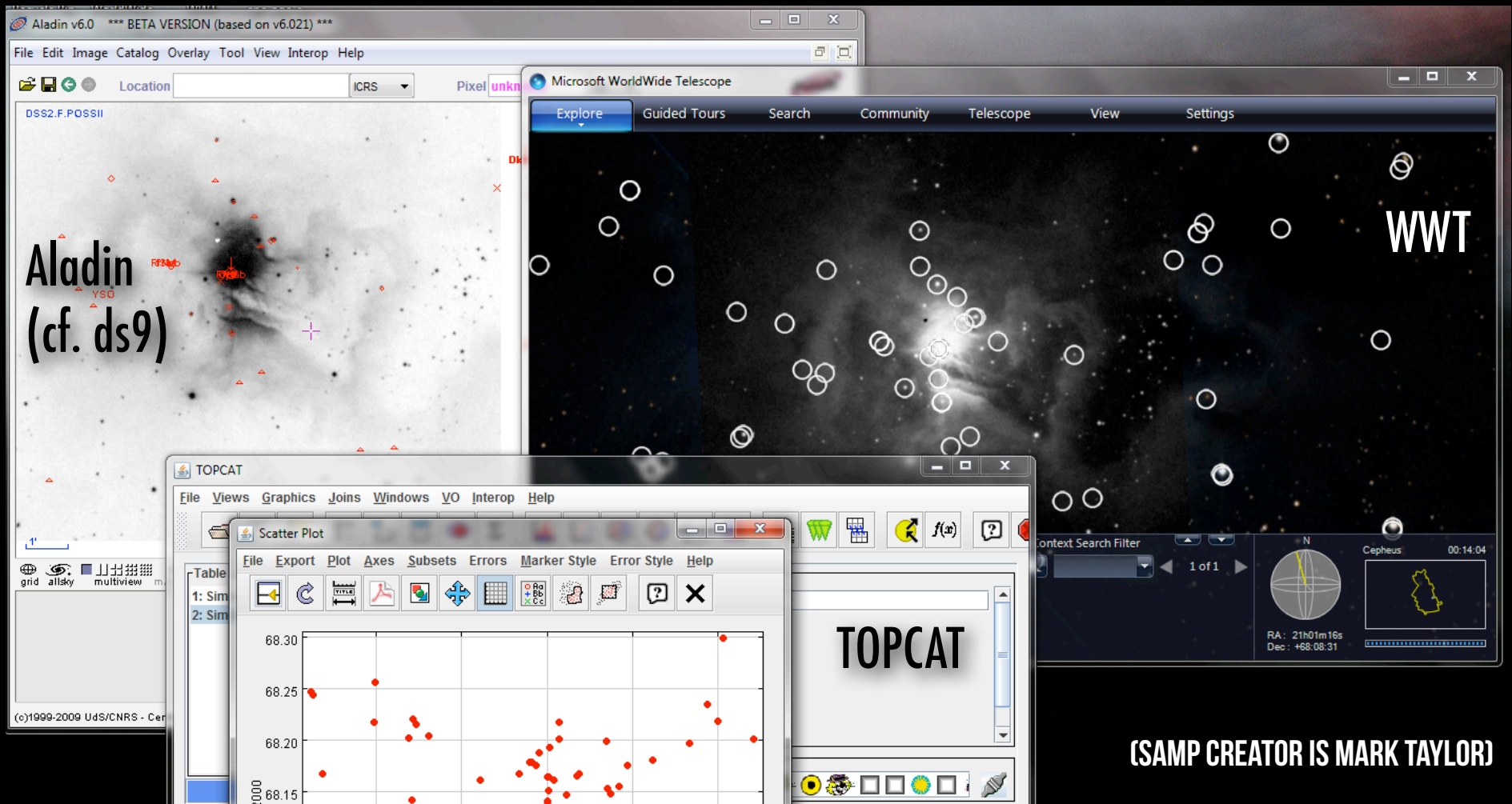


John Tukey



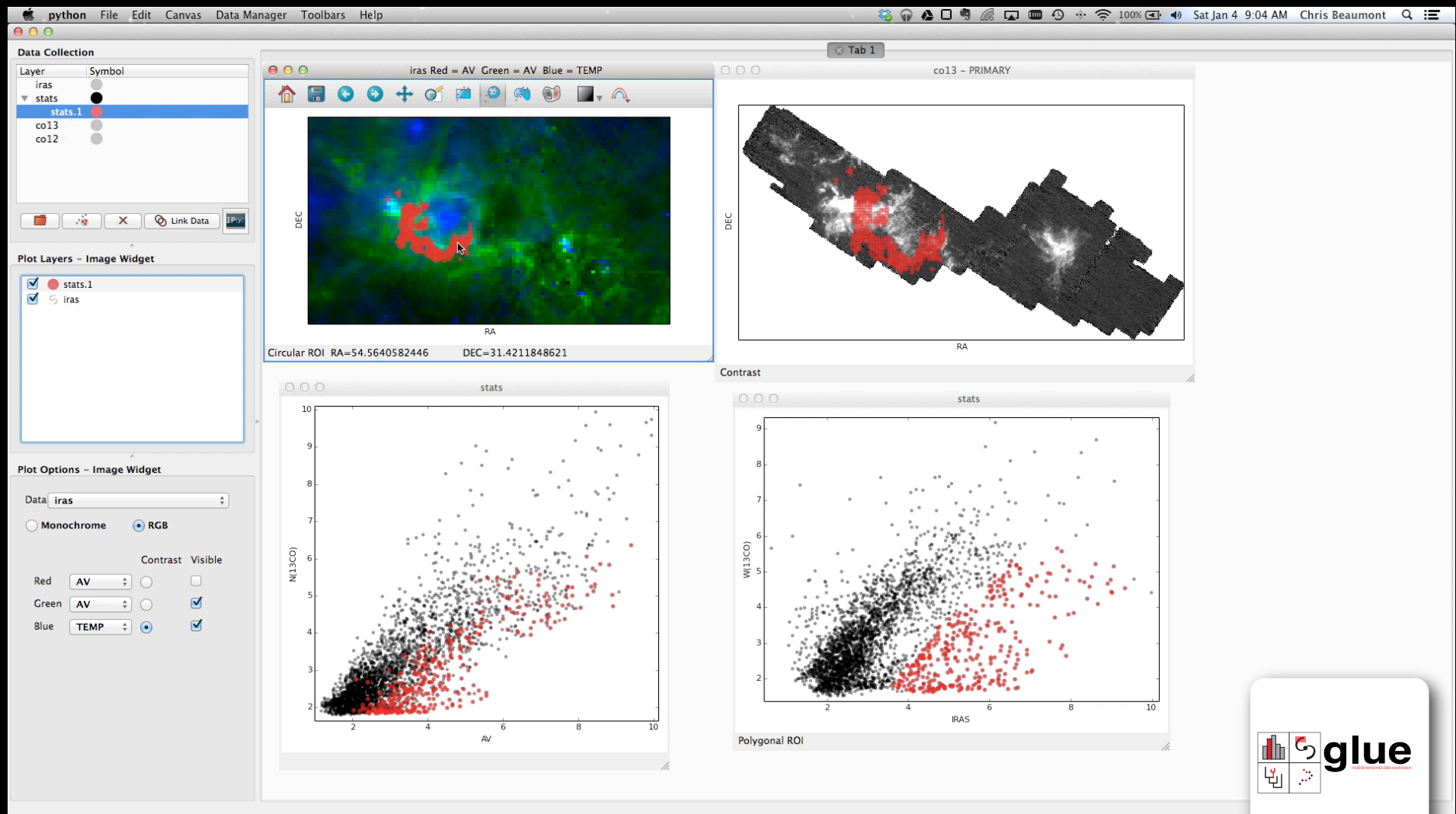
figure, by M. Borkin, reproduced from [Goodman 2012](#), "Principles of High-Dimensional Data Visualization in Astronomy"

# LINKED VIEWS OF HIGH-DIMENSIONAL DATA "SAMP"



figure, showing SAMP screenshot, reproduced from Goodman 2012, "Principles of High-Dimensional Data Visualization in Astronomy"

# LINKED VIEWS OF HIGH-DIMENSIONAL DATA GLUE





1610



SIDEREUS NUNCIIUS

On the third, at the seventh hour, the sequence. The eastern one was 1 minute, the closest western one 2 minutes, and the

East \* ○ \* 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 \* ○ \* \* 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 1/2 than the rest. But at the seventh hour the eastern 1/30 seconds apart, Jupiter was a minutes from the

East \*\* ○ \*

one, while he was 4 minutes from the next western one was 1 minutes from the westernmost one. They and extended on the same straight line along the celestial

On the fifth, the sky was cloudy.

On the sixth, only two stars appeared flanking Jupiter

East \* ○ \*

in the adjoining figure. The eastern one was 2 minutes western one 1 minutes from Jupiter. They were on the line with Jupiter and equal in magnitude.

On the seventh, two stars stood near Jupiter, but arranged in this manner.

1665



1895

ASTROPHYSICAL JOURNAL

AN INTERNATIONAL REVIEW OF SPECTROSCOPY AND ASTRONOMICAL PHYSICS

VOLUME I JANUARY 1895 NUMBER 1

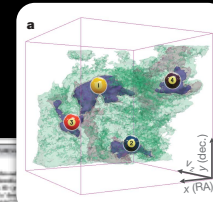
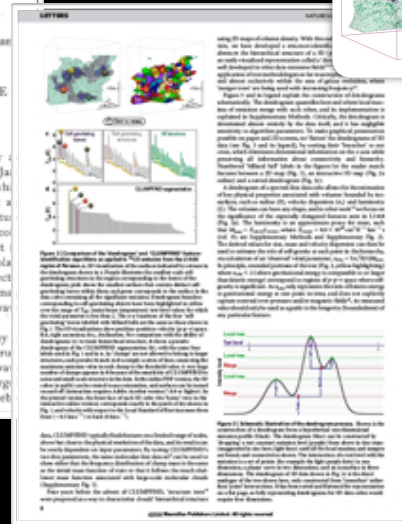
ON THE CONDITIONS WHICH AFFECT THE SPECTRO-PHOTOGRAPHY OF THE SUN.

By ALBERT A. MICHELSON.

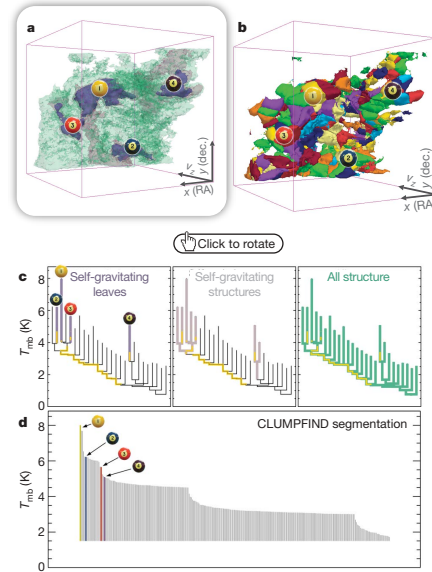
The recent developments in solar spectro-photography in great measure due to the device originally suggested by Jansen and perfected by Hale and Deslandres, by means of which a photograph of the Sun's prominences may be obtained at a time as readily as it is during an eclipse. The essential feature of this device are the simultaneous movements of the camera-rotator across the Sun's image, with that of a second slit (the focus of the photographic lens) over a photographic plate. If these relative motions are so adjusted that the same spectral line always falls on the second slit, then a photographic image of the Sun will be reproduced by light of this particular wavelength.

Evidently the process is not limited to the photography of the prominences, but extends to all other peculiarities of structure which emit radiations of approximately constant wavelength; and the efficiency of the method depends very largely upon the contrast which can be obtained by the greater extent

2009



# WHAT DO WE PUBLISH?



**Figure 2 | Comparison of the 'dendrogram' and 'CLUMPFIND' feature-identification algorithms as applied to  $^{13}\text{CO}$  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  $T_{\text{mb}}$  (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 ( $-0.5 \text{ km s}^{-1}$ ) to back ( $8 \text{ km s}^{-1}$ ).

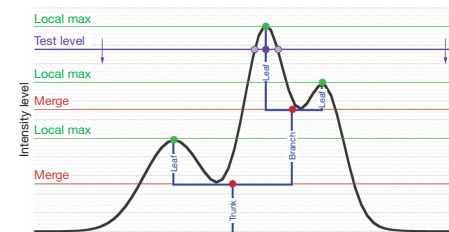
data, CLUMPFIND typically finds features on a limited range of scales, above but close to the physical resolution of the data, and its results can be overly dependent on input parameters. By tuning CLUMPFIND's two free parameters, the same molecular-line data set<sup>8</sup> can be used to show either 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'<sup>9</sup> were proposed as a way to characterize clouds' hierarchical structure

using 2D maps of column density. With this work as inspiration, we have developed a structure-identification algorithm that abstracts the hierarchical structure of a data cube into an easily visualized representation called a dendrogram. Well developed in other data-intensive applications, the application of tree methodologies so far has been almost exclusively within the astronomical domain. 'merger trees' are being used with increasing frequency.

Figure 3 and its legend explain the dendrogram process schematically. The dendrogram quantifies the hierarchical structure of the data cube with each node of emission merge with each other. The dendrogram is explained in Supplementary Methods and Supplementary Fig. 2. The sensitivity to algorithm parameters is discussed in the Supplementary Information on paper and 2D screen data (see Fig. 3 and its legend). The dendrogram is a 2D representation of a 3D data cube, preserving all information. The dendrogram is a 2D representation of a 3D data cube, preserving all information. The dendrogram is a 2D representation of a 3D data cube, preserving all information.

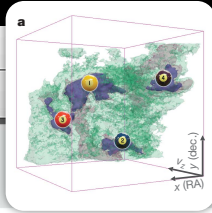
A dendrogram of a spectral line emission cube, such as radius ( $R$ ), luminosity ( $L$ ). The volumes can have any shape, and are defined by the significance of the especially elongated features (Fig. 2a). The luminosity is an approximate proxy for mass,  $M_{\text{lum}} = X_{13\text{CO}} L_{13\text{CO}}$ , where  $X_{13\text{CO}} = 8.0 \times 10^{20} \text{ cm}^{-2} \text{ K}^{-1} \text{ km}^{-1} \text{ s}$  (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,  $\alpha_{\text{obs}} = 5\sigma_v^2 R/GM_{\text{lum}}$ . In principle, extended portions of the tree (Fig. 2, yellow highlighting) where  $\alpha_{\text{obs}} < 2$  (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  $\alpha_{\text{obs}}$  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

2009  
3D PDF  
INTERACTIVITY  
IN A "PAPER"



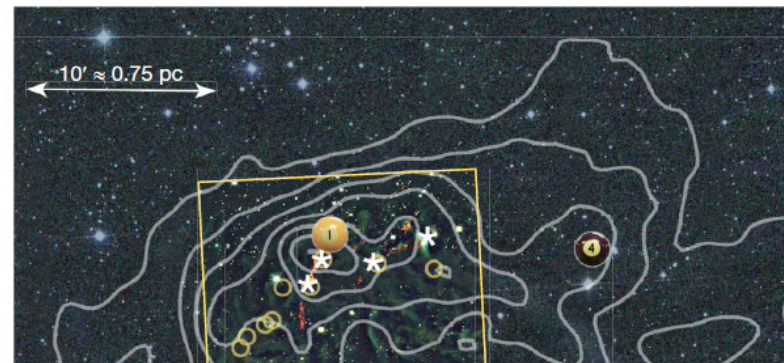
# LETTERS

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

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>

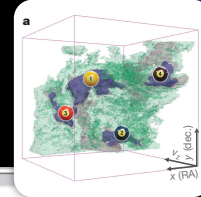
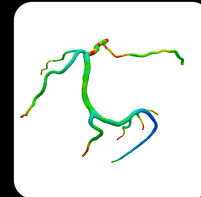
Self-gravity plays a decisive role in the final stages of star formation, where dense cores (size  $\sim 0.1$  parsecs) inside molecular clouds collapse to form star-plus-disk systems<sup>1</sup>. But self-gravity's role at earlier times (and on larger length scales, such as  $\sim 1$  parsec) is unclear; 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<sup>2</sup>. 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 emission<sup>3</sup> 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



# TOWARD THE PAPER OF THE FUTURE

2009



2014



**Beyond Galileo**  
Josh Peek, Alberto Pepe, Adrian Price-Whelan, Chris Beaumont

In the last portion of *Sidereus Nuncius*, Galileo reported his discovery of **four objects** that appeared to form a straight line of stars near **Jupiter**. The first night, he witnessed a line of three title stars close to Jupiter parallel to the ecliptic; the following nights brought different arrangements and another star into his view, totaling four stars around Jupiter. (Galilei 1610) Throughout the text, Galileo gave illustrations of the relative positions of Jupiter and its apparent companion stars as they appeared nightly from late January through early March 1610. The fact that they changed their positions relative to Jupiter from night to night, but always appeared in the same straight line near Jupiter, brought Galileo to deduce that they were four bodies in orbit around Jupiter. On January 11 after 4 nights of observation he wrote:

"I therefore concluded and decided unhesitatingly, that there are three stars in the heavens moving about Jupiter, as Venus and Mercury round the Sun; which at length was established as clear as daylight by numerous subsequent observations. These observations also established that there are not only three, but four, errant sidereal bodies performing their revolutions round Jupiter...the revolutions are so swift that an observer may generally get differences of position every hour." (Galilei 1610)

**Inclination (deg) vs Eccentricity**

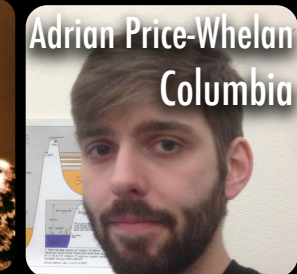
Eccentricity	Inclination (deg)
0.0	0.0
0.1	0.0
0.2	0.0
0.3	0.0
0.4	0.0
0.5	0.0
0.6	0.0

**log Mass (log kg) vs Semi-major axis (km)**

Semi-major axis (km)	log Mass (log kg)
0	14
1e+7	14
2e+7	14
3e+7	14

Four Centuries of Discovery | A Chasm in Mass | Some are Similar... | ...but Most are Different

# "THE STORY & THE SANDBOX" (GLUE:D3PO:AUTHOREA)



https://www.authorea.com/users/2786/articles/4039/\_show\_article

Authorea

BROWSE ABOUT CONTACT PLANS FEEDBACK HELP JOSH PEEK ▾



Index ROUGH DRAFT OPEN SCIENCE

Settings

Fork

Quick edit ▾

Tour

0 Comments

Export

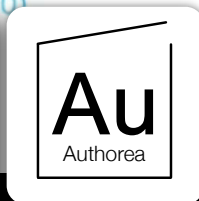


## Beyond Galileo

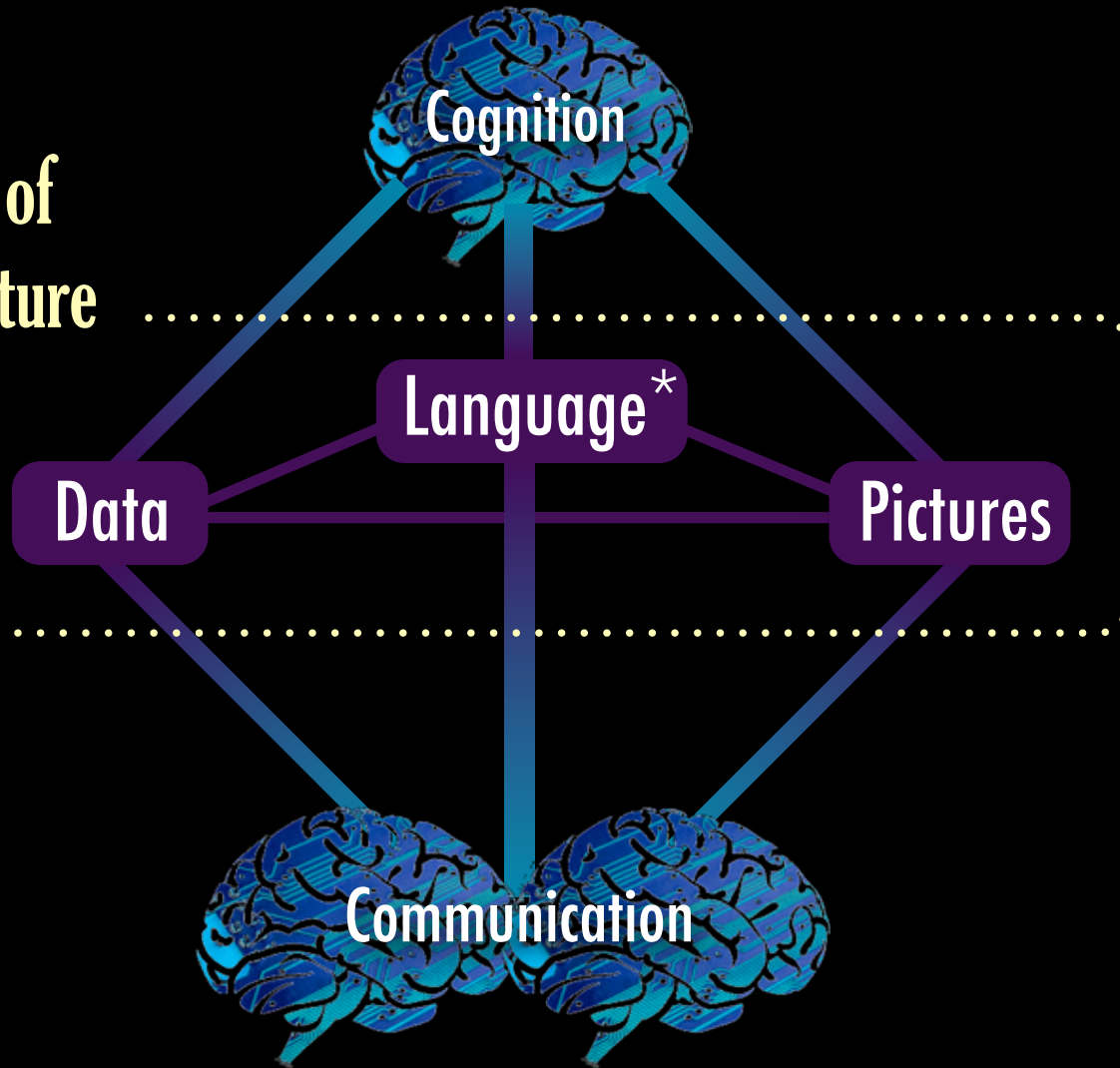
Josh Peek, Alberto Pepe [+ Add author](#)

In the last portion of *Sidereus Nuncius*, Galileo reported his discovery of **four objects** that appeared to form a straight line of stars near **Jupiter**. The first night, he witnessed a line of three little stars close to Jupiter parallel to the ecliptic; the following nights brought different arrangements and another star into his view, totaling four stars around Jupiter. (Galilei 1618) Throughout the text, Galileo gave illustrations of the relative positions of Jupiter and its apparent companion stars as they appeared nightly from late January through early March 1610. The fact that they changed their positions relative to Jupiter from night to night, but always appeared in the same straight line near Jupiter, brought Galileo to deduce that they were four bodies in orbit around Jupiter. On January 11 after 4 nights of observation he wrote:

"I therefore concluded and decided unhesitatingly, that there are three stars in the heavens moving about Jupiter, as Venus and Mercury round the Sun; which at length was established as clear as daylight by numerous subsequent observations. These observations also established that there are not only three, but four, erratic sidereal bodies performing their revolutions round Jupiter...the revolutions are so swift that an observer may generally get differences of position every hour." (Galilei 1610)



# Paper of the Future



\*"Language" includes words & math



# The "Paper" of the Future

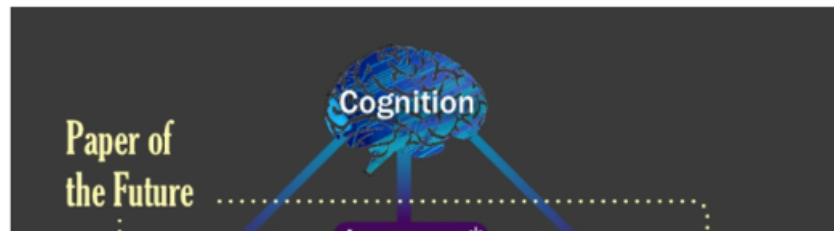
Alyssa Goodman, Josh Peek, Alberto Accomazzi, Chris Beaumont, Christine L. Borgman, How-Huan Hope Chen, Merce Crosas, Christopher Erdmann, August Muench, Alberto Pepe, Curtis Wong

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 communicate 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.**



3

2

0

Konrad Hinsien 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 will...

[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 qua...

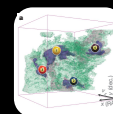
[more](#)

Konrad Hinsien 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.

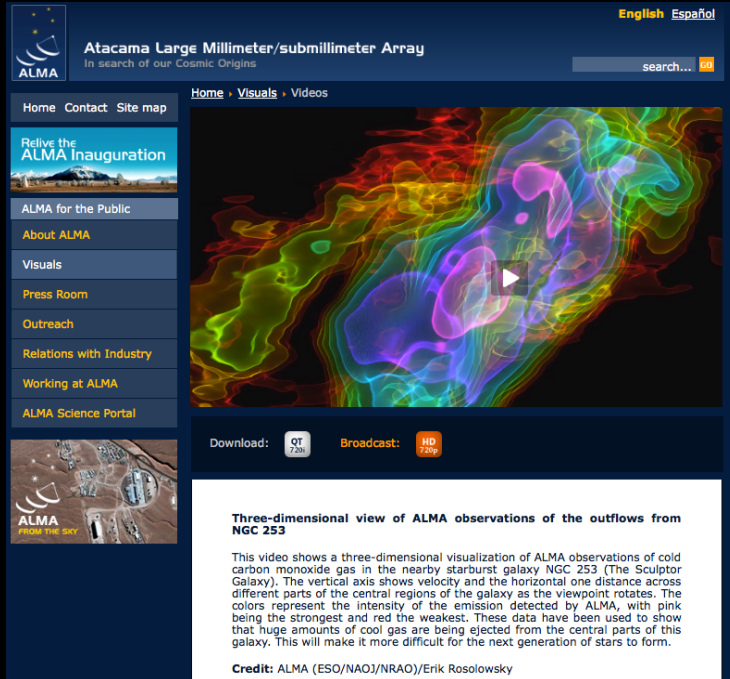
Christine L. Borgman 4 months ago · Private

"publications"



[video, demos]

# THE FUTURE IS IN 3D



English Español

Atacama Large Millimeter/submillimeter Array  
In search of our Cosmic Origins

Home Contact Site map Home Visuals Videos

Relive the ALMA Inauguration

ALMA for the Public

About ALMA

Visuals




Press Room

Outreach

Relations with Industry

Working at ALMA

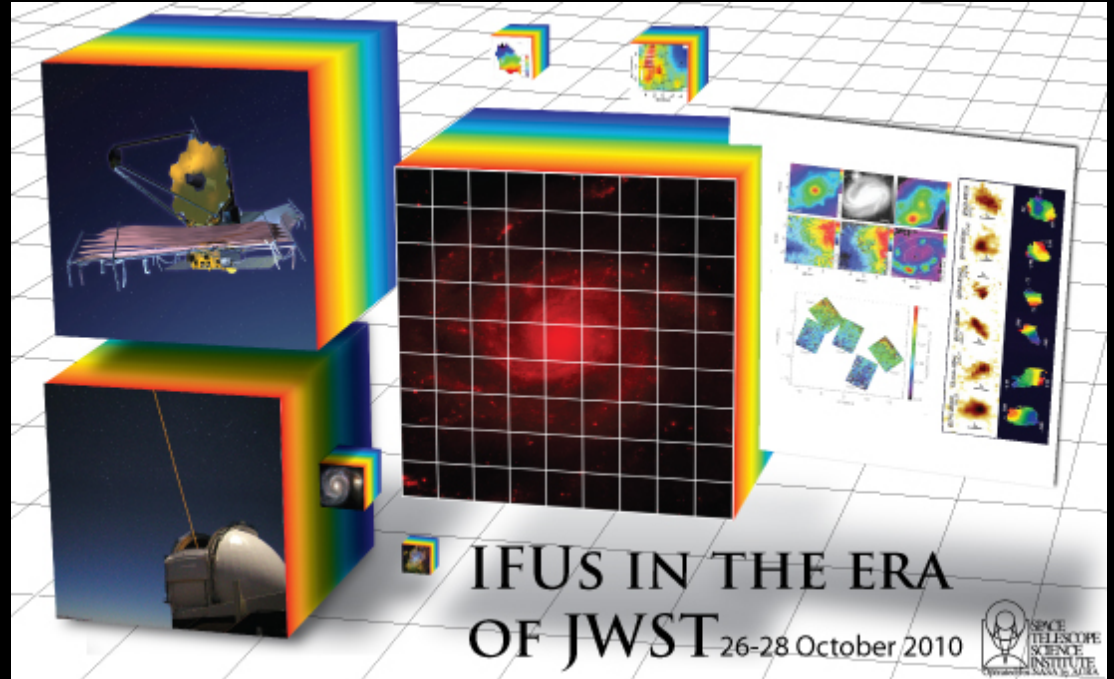
ALMA Science Portal

Download:   Broadcast: 


**Three-dimensional view of ALMA observations of the outflows from NGC 253**

This video shows a three-dimensional visualization of ALMA observations of cold carbon monoxide gas in the nearby starburst galaxy NGC 253 (The Sculptor Galaxy). The vertical axis shows velocity and the horizontal one distance across different parts of the central regions of the galaxy as the viewpoint rotates. The colors represent the intensity of the emission detected by ALMA, with pink being the strongest and red the weakest. These data have been used to show that huge amounts of cool gas are being ejected from the central parts of this galaxy. This will make it more difficult for the next generation of stars to form.

Credit: ALMA (ESO/NAO/NRAO)/Erik Rosolowsky



IFUS IN THE ERA  
OF JWST 26-28 October 2010



yt viz from ALMA data  
(Turk, Rosolowsky)

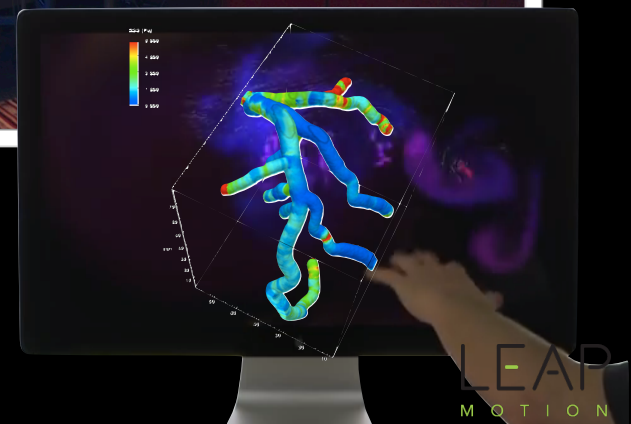
IFUs on JWST...with Glue!  
(Glue funded by NASA)



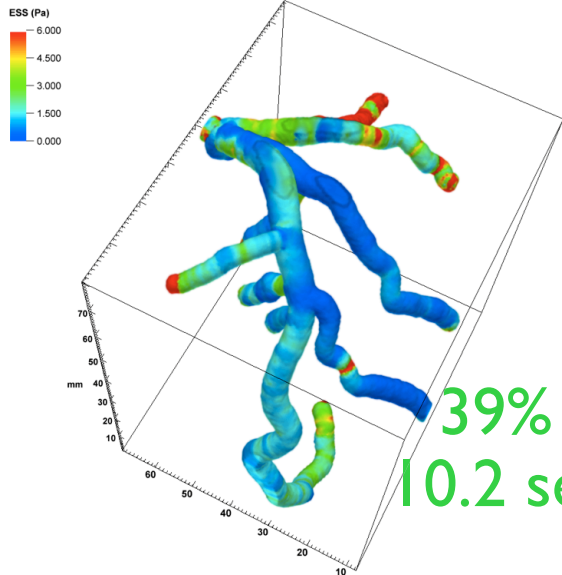
# THE FUTURE IS MODULAR, OPEN-SOURCE, AND NOT (JUST) ON THE DESKTOP



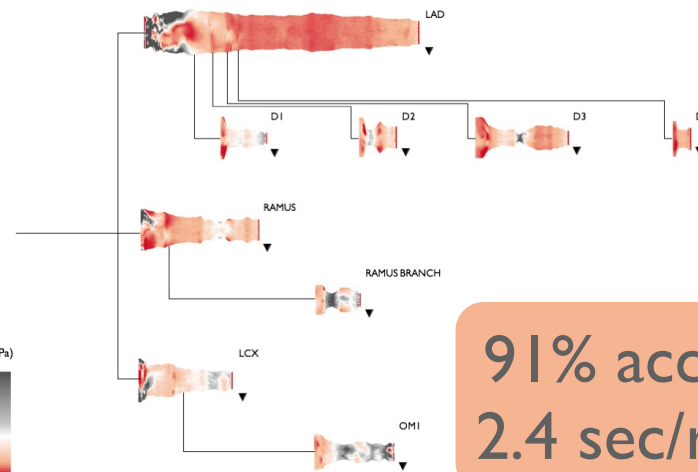
**HACK  
TO THE  
FUTURE**



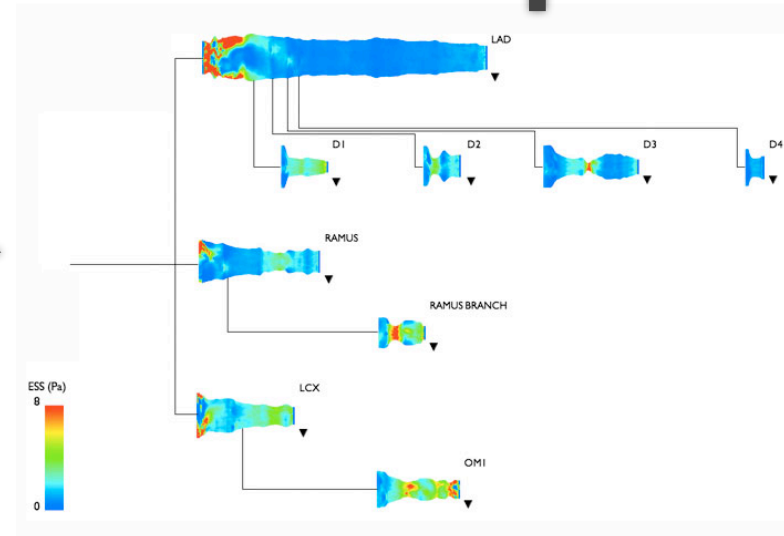
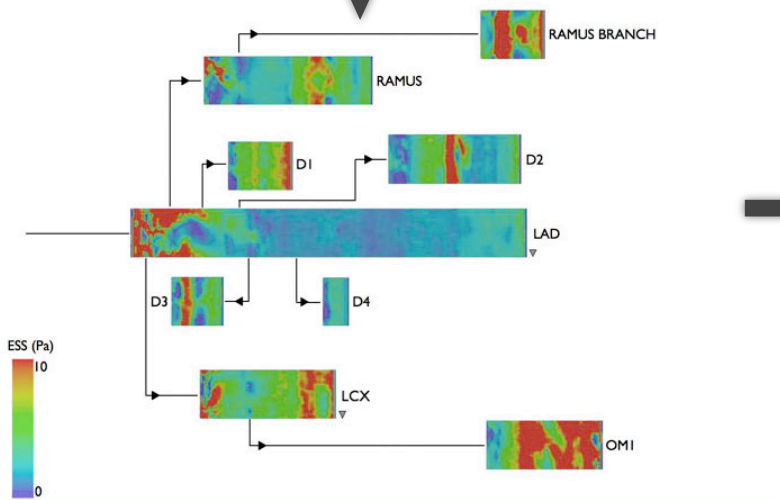
# DIMENSIONALITY AND COLOR



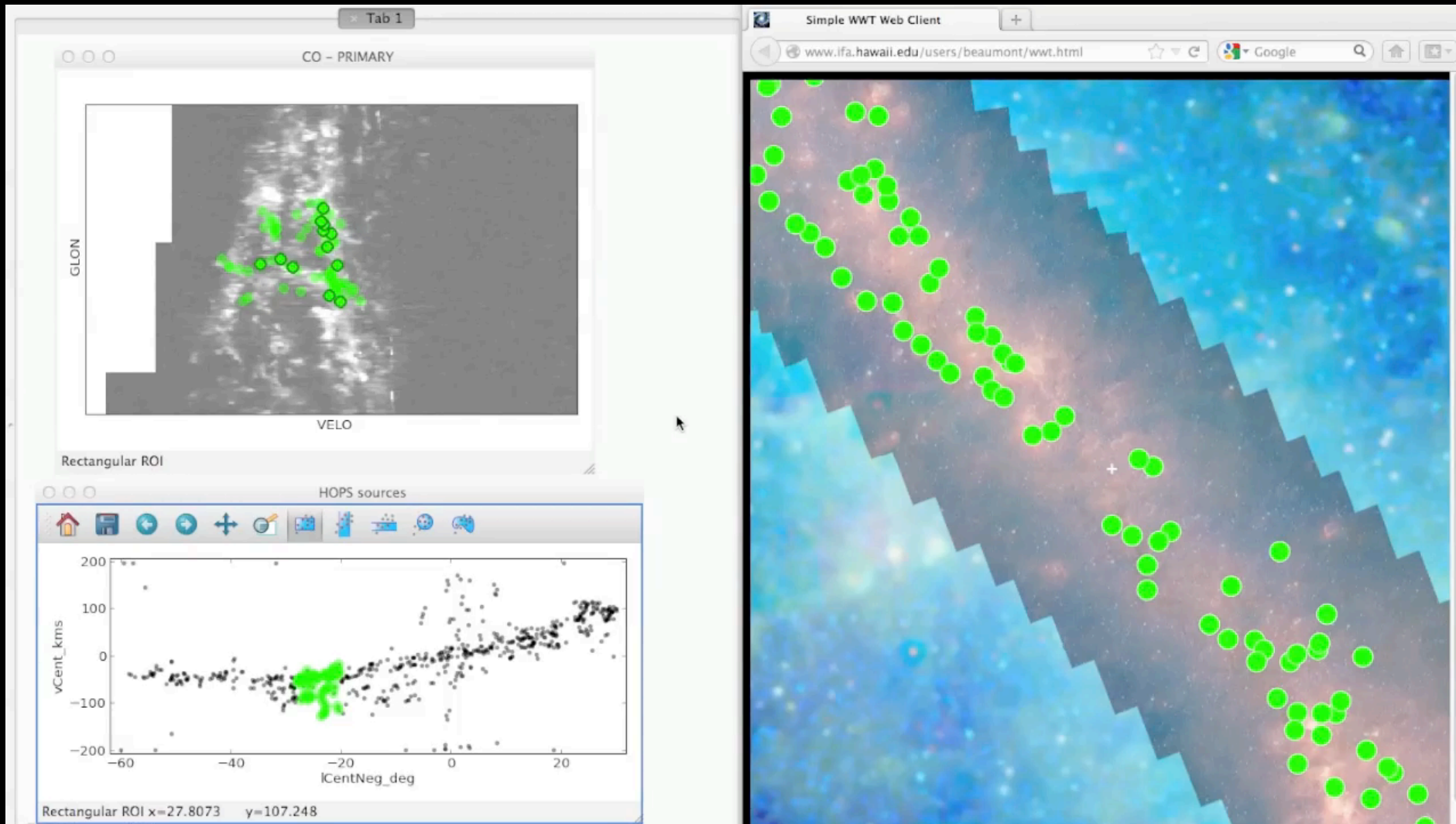
39% accurate  
10.2 sec/region



91% accurate  
2.4 sec/region



Borkin et al. 2011  
cf. colorbrewer2.org



Video courtesy of Chris Beaumont, Lead Glue Architect

# THE FUTURE OFFERS NEW WAYS TO LEARN

## WorldWide Telescope Ambassadors



## Higher Ed

the 2013 experiment

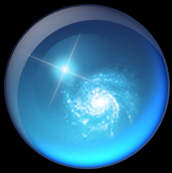
HARVARD UNIVERSITY  
ASTRONOMY 201B  
DEMOFEST

LOCATION  
Perkin Lobby and Wolbach Library, 60 Garden Street

TIME  
11-12 for drop-in demos  
12-12:45 lunch for students & their guests

PREVIEW  
<http://ay201b.wordpress.com/topical-modules>

The complex block contains a vertical text element on the left, a main title, logos for ISM, edX, and VERI TAS, location and time information, and a preview link. The background features a starry space theme.



# Microsoft® Research WorldWide Telescope

[worldwidetelescope.org](http://worldwidetelescope.org)

The screenshot shows the WorldWide Telescope interface with several key components:

- Navigation Bar:** Includes tabs for 'Explore', 'Guided Tours', 'Search', 'View', and 'Settings'.
- Collections:** A row of thumbnails for different astronomical surveys, including 'Digitized Sky Survey', 'VLSS: VLA Low-frequency Sky Survey', 'WMAP ILC 5-Year Temperature Anisotropy Map', 'SFD Dust Map (Infrared)', 'IRIS: Improved Resolution', '2MASS: Two Micron All Sky Survey', and 'Hydrogen Alpha Filter'.
- Main View:** A large central window displaying a 3D view of the Andromeda Galaxy (M31) with a 'Finder Scope' overlaid.
- Finder Scope Panel:** Provides detailed information for the selected object, NGC224, including its classification as a 'Spiral Galaxy in Andromeda', RA (00h42m42s), Dec (41:16:00), Distance, Magnitude, Alt (70:06:26), Rise, Az (275:42:17), and Transit (00:35). It also includes 'Research' and 'Show Object' buttons.
- Context Bar:** Located at the bottom, it shows a 'Look At' dropdown set to 'Sky', a 'Context globe' showing the current field of view, and a 'Context bar' with thumbnails for 'NGC221' and 'M31'.
- Image Credits:** A panel at the bottom left provides information about the data sources, mentioning NASA satellites IRAS and COBE.

Seamlessly explore imagery from the best ground and space-based telescopes in the world

Expert led tours of the Universe

Control time to study how the night sky changes

View and compare images from across the electromagnetic spectrum

Much more than "just" the sky at night! 3D features can take you to other planets, stars & galaxies.

Finder Scope links to Wikipedia, publications, and data, so you can learn more

Context bar shows items of interest in current field of view

Context globe shows where you're looking.



# WHAT DOES "Shells in Perseus" LOOK LIKE IN THE FUTURE?

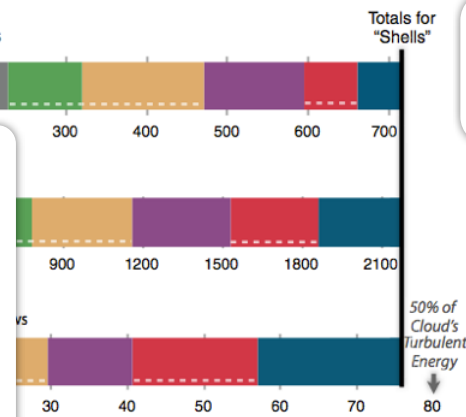
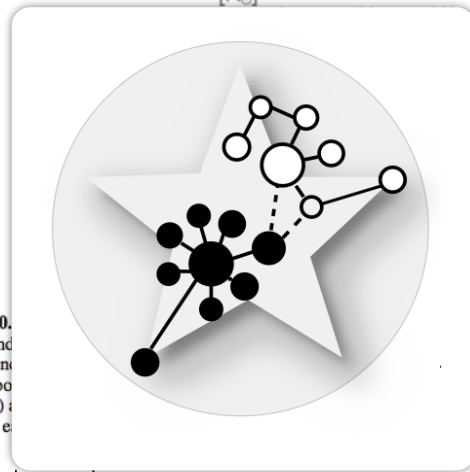
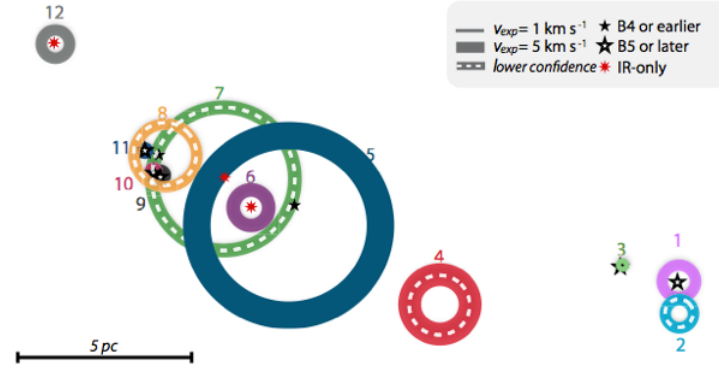


Figure 30. radius (and a confident star symbol known  $\alpha$ ) value for e

tribution of shells in Perseus. The radius of each circle (and position) are proportional to the expansion velocity (see the legend on the upper right corner). Shells with Candidate powering sources with a B5 spectral type or later are shown as white (filled) star symbols. Candidate sources with no known spectral type (but energy of the shells are shown in the three horizontal bars (where the colors indicate the energy of the molecular outflows in Perseus (from Arce et al. 2010) are shown for comparison.

THE ASTROPHYSICAL JOURNAL, 742:105 (30pp), 2011 December 1

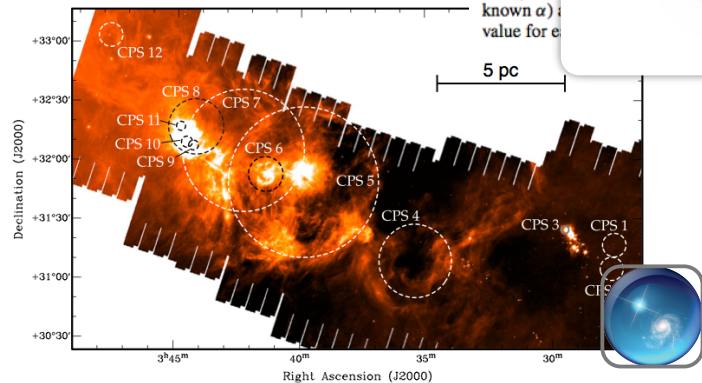
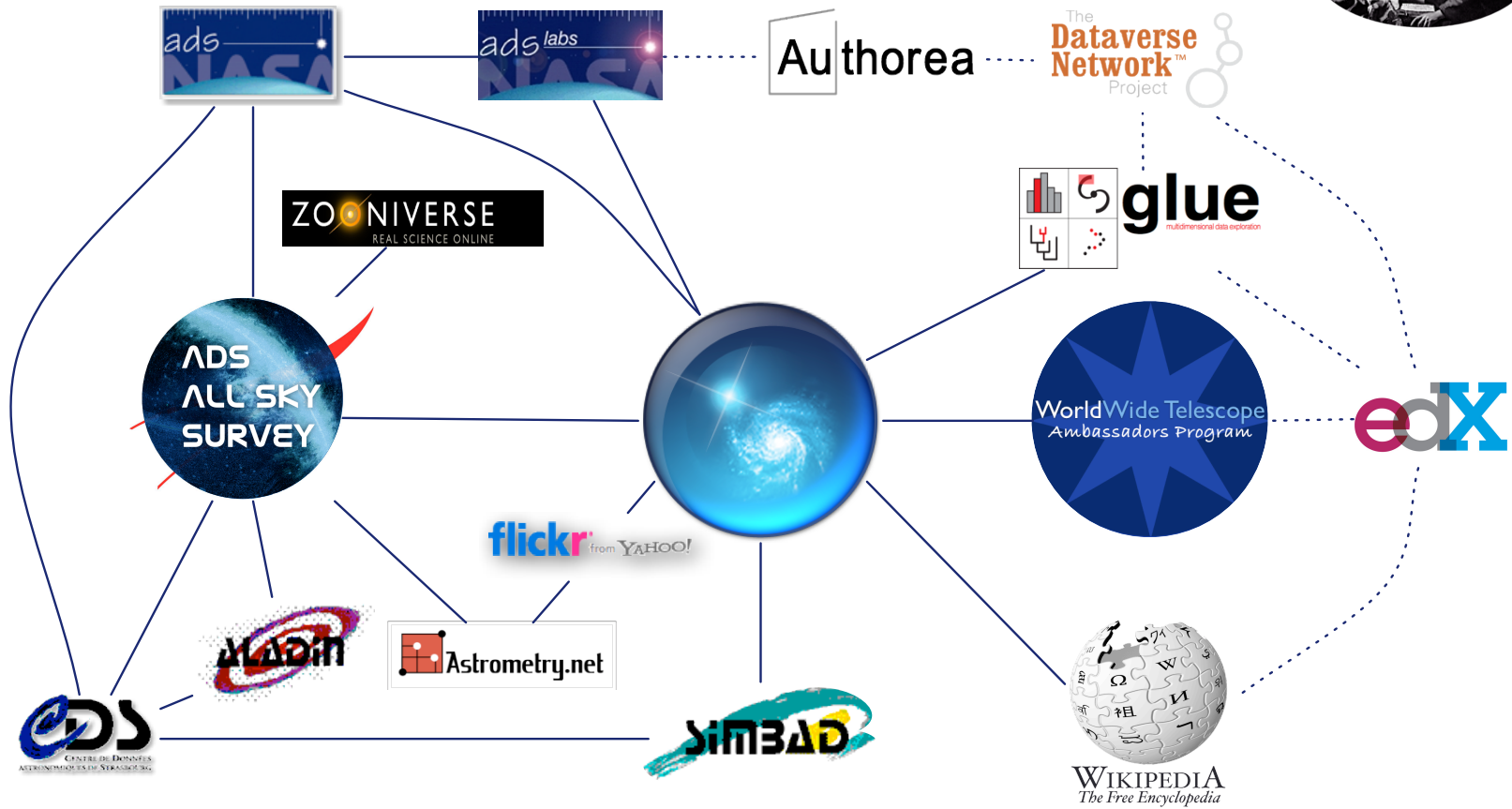


Figure 1. Shells in Perseus. The position of each CPS is shown as a dashed circle overlaid on the c2d *Spitzer* MIPS 1 ( $24 \mu\text{m}$ ) image (from Rebull et al. 2007). (A color version of this figure is available in the online journal.)



# SEAMLESS ASTRONOMY

Linking scientific data, publications, and communities



<https://www.cfa.harvard.edu/~agoodman/seamless/>

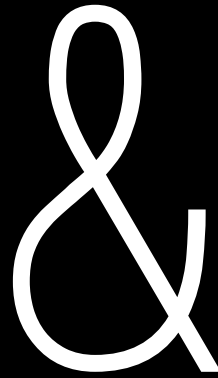
Supported by

Microsoft Research



Made possible by MANY collaborators, listed at [projects.iq.harvard.edu/seamlessastronomy](http://projects.iq.harvard.edu/seamlessastronomy)

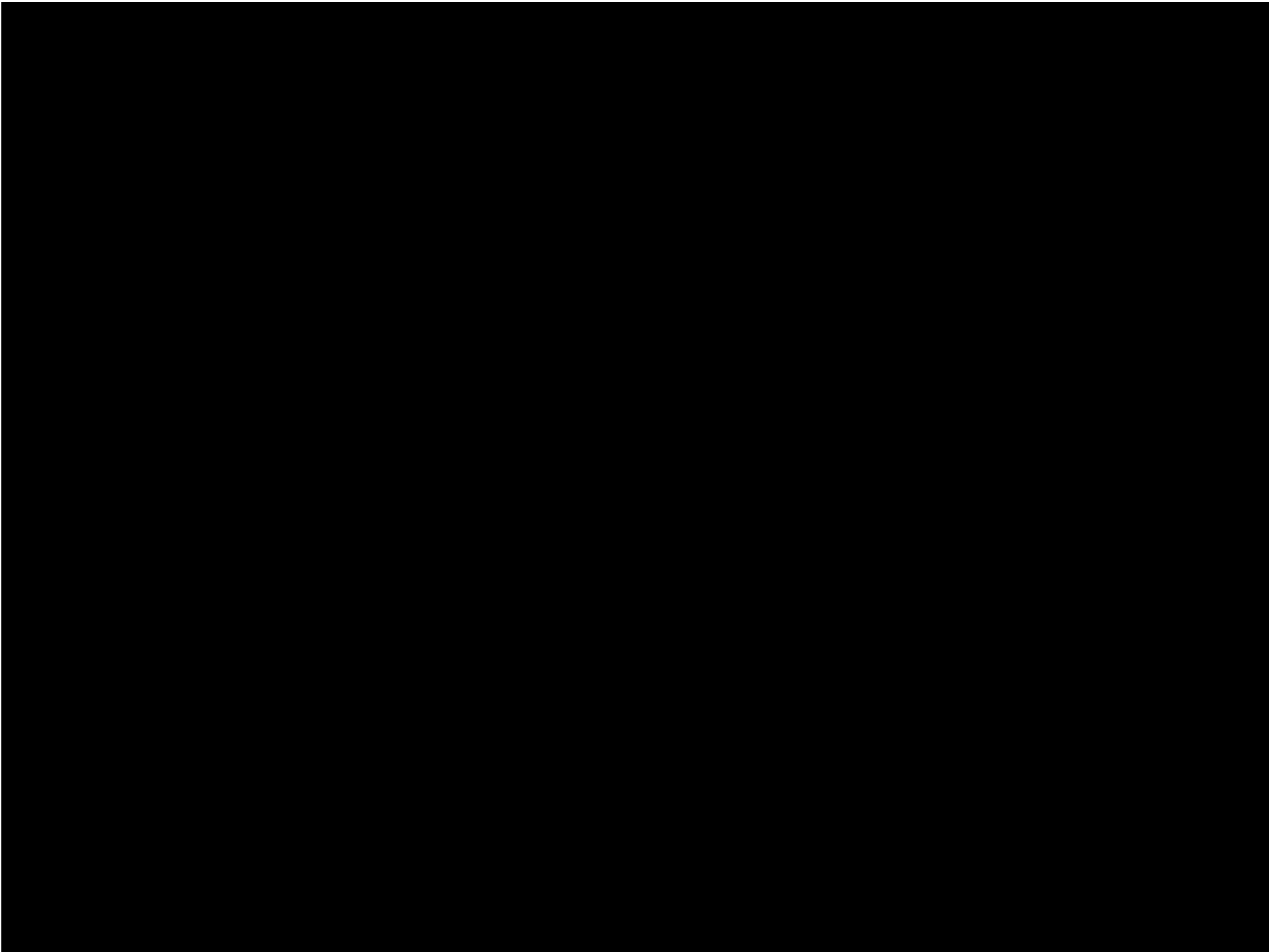
HIGH-DIMENSIONAL  
DATA  
VISUALIZATION

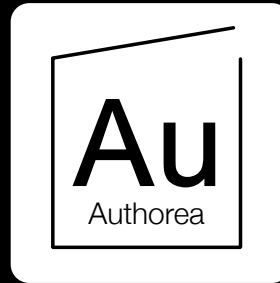
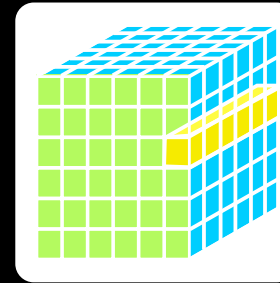
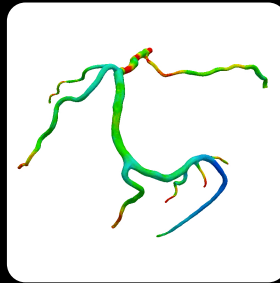
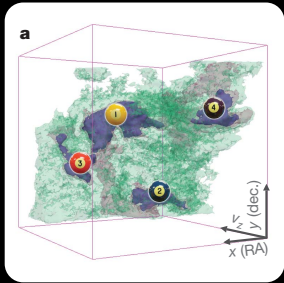
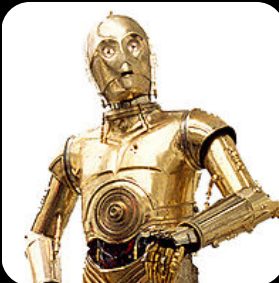
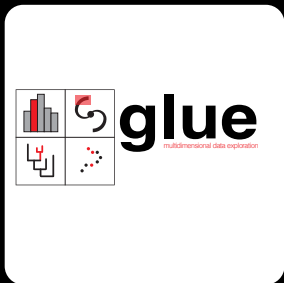
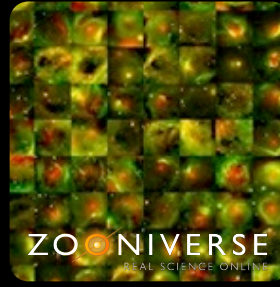
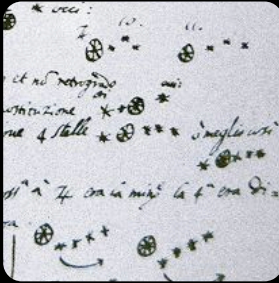


THE "PAPER"  
OF THE  
FUTURE

**Alyssa A. Goodman • Harvard-Smithsonian Center for Astrophysics**







JOIN US?

