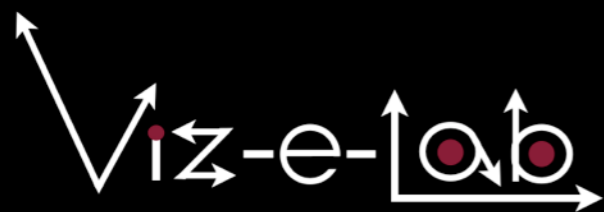
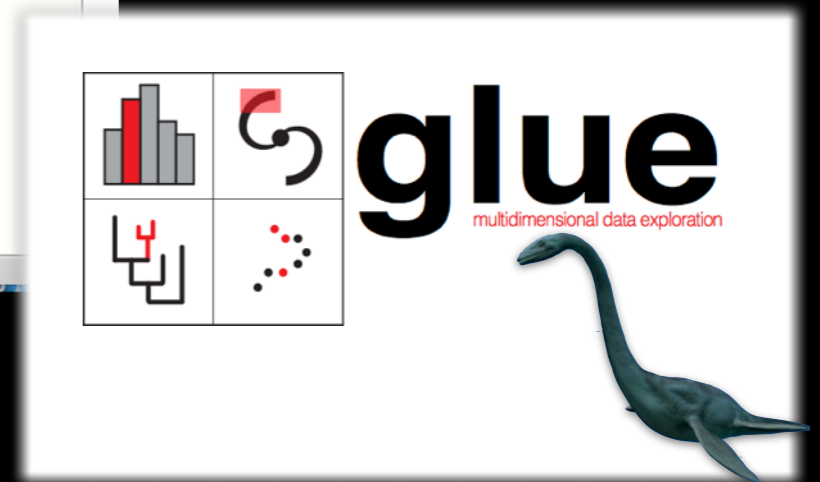
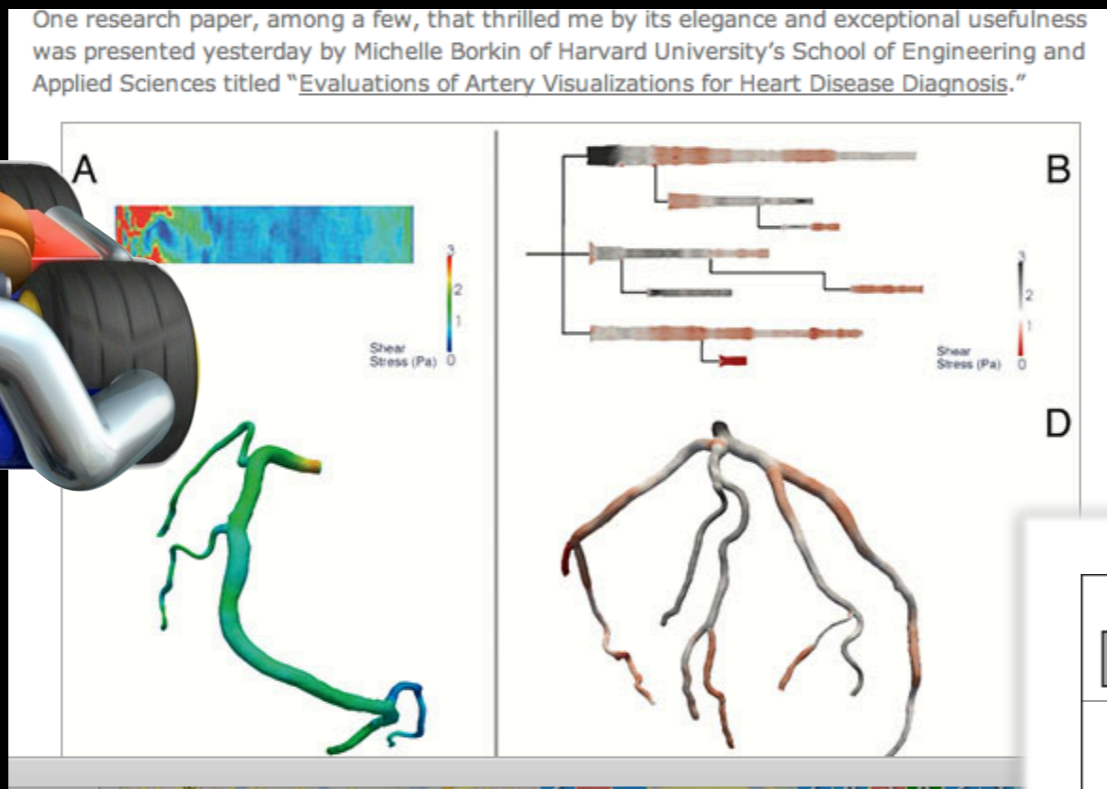


Seeing More in Data



Alyssa A. Goodman
Harvard-Smithsonian Center for Astrophysics

Seeing More in Data

SEAMLESS ASTRONOMY
Linking scientific data, publications, and communities

ABOUT PROJECTS PEOPLE RESOURCES DATAVERSE

SEAMLESS ASTRONOMY

About

Any and all paths to data

Virtual Observatory tools of variable utility

Advanced Search & InfoViz tools

Standalone Analysis Software

Almost exclusively ADS

Google

ADS

nature

arXiv.org

The Seamless Astronomy Group at the Harvard-Smithsonian Center for Astrophysics brings together astronomers, computer scientists, information scientists, librarians and visualization experts involved in the development of tools and systems to study and enable the next generation of online astronomical research.

Current projects include research on the development of systems that seamlessly integrate scientific data and literature, the semantic interlinking and annotation of scientific resources, the study of the impact of social media and networking sites on scientific dissemination, and the analysis and visualization of astronomical research communities. Visit our [project page](#) to find out more.

SHARE [social media icons]

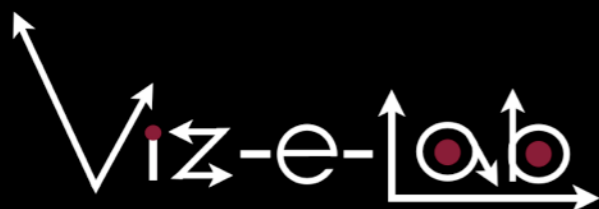
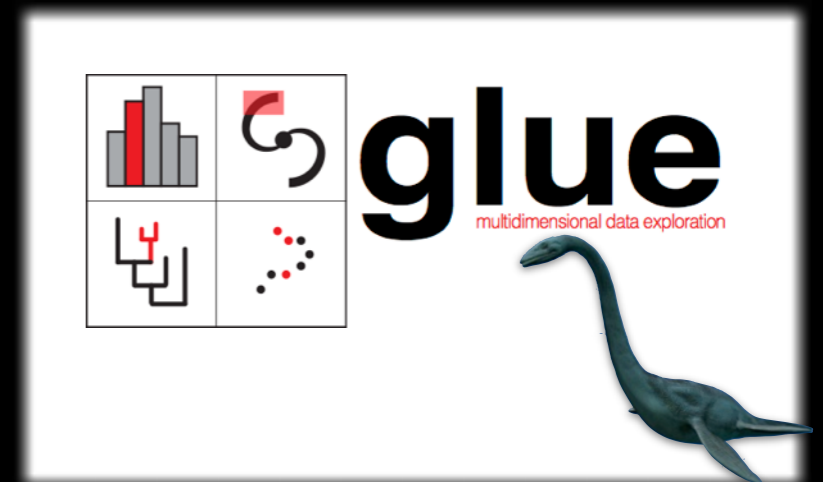
Latest Announcements

Introducing the Astronomy Dataverse

Latest Feed Items

@rahuldave there is a writeboard with my notes... More at next #seamlessastronomy next week.

Thanks to @astrobites and @astroknight06 for great summary <http://t.co/jWWFT0CD> of our High-D Data Viz work! #ivoa #seamlessastronomy



Alyssa A. Goodman
Harvard-Smithsonian Center for Astrophysics

Relative Strengths



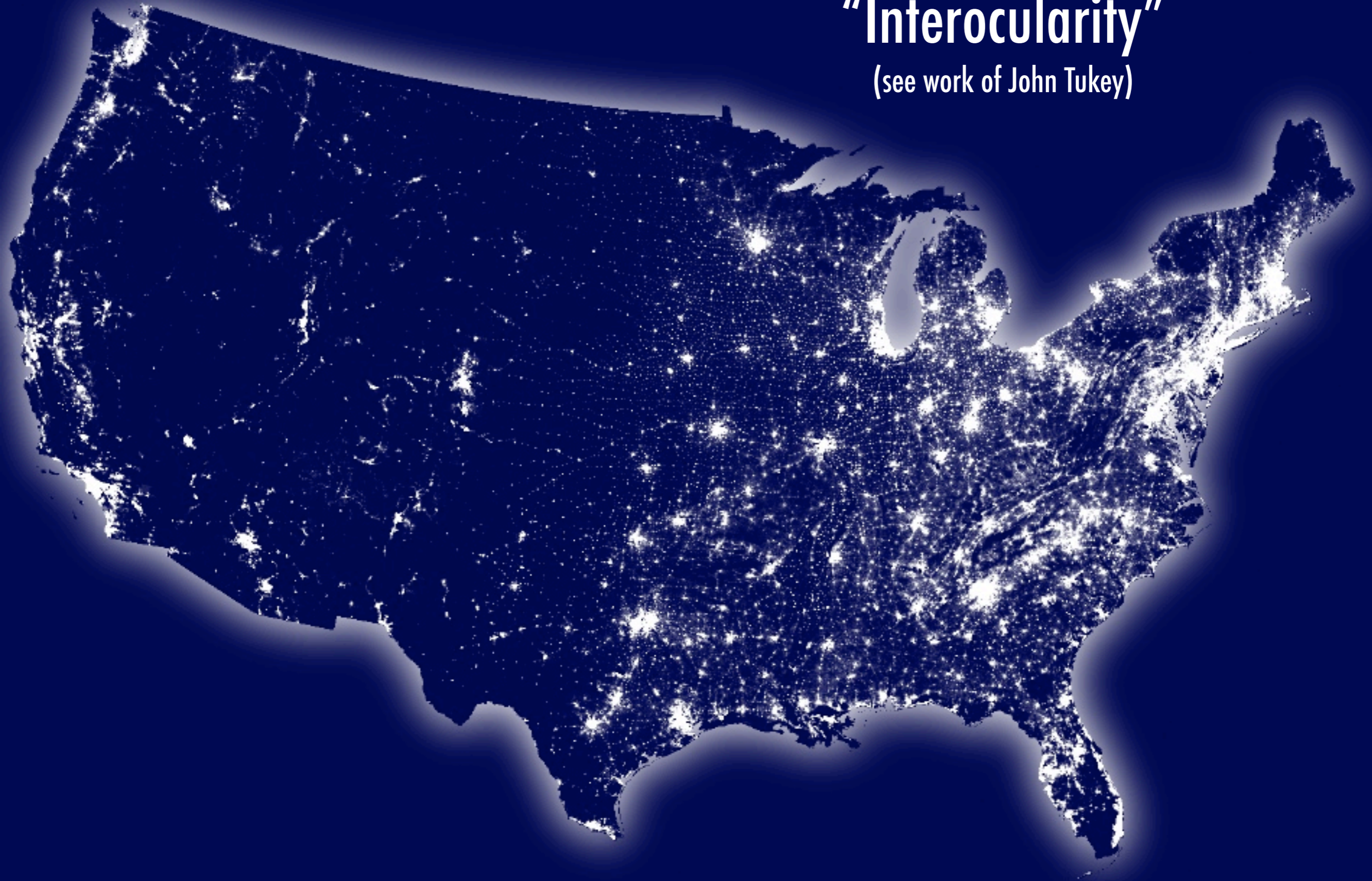
Pattern Recognition
Creativity



Calculations

“Interocularity”

(see work of John Tukey)



What...

...is easier now than before?

fast computation, animation, 3D

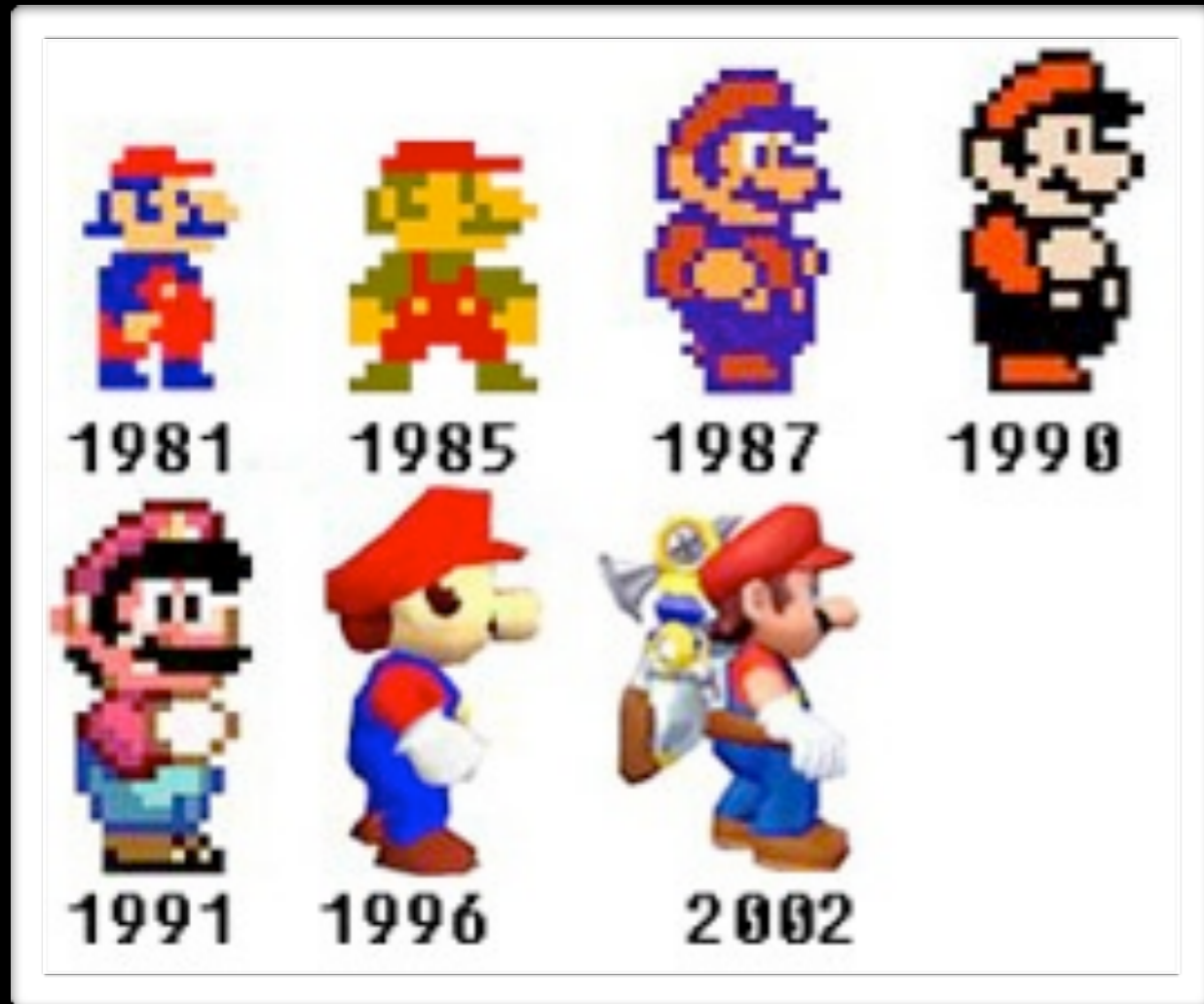
...was easier before than now?

craftsmanship

...should be easier in the future?

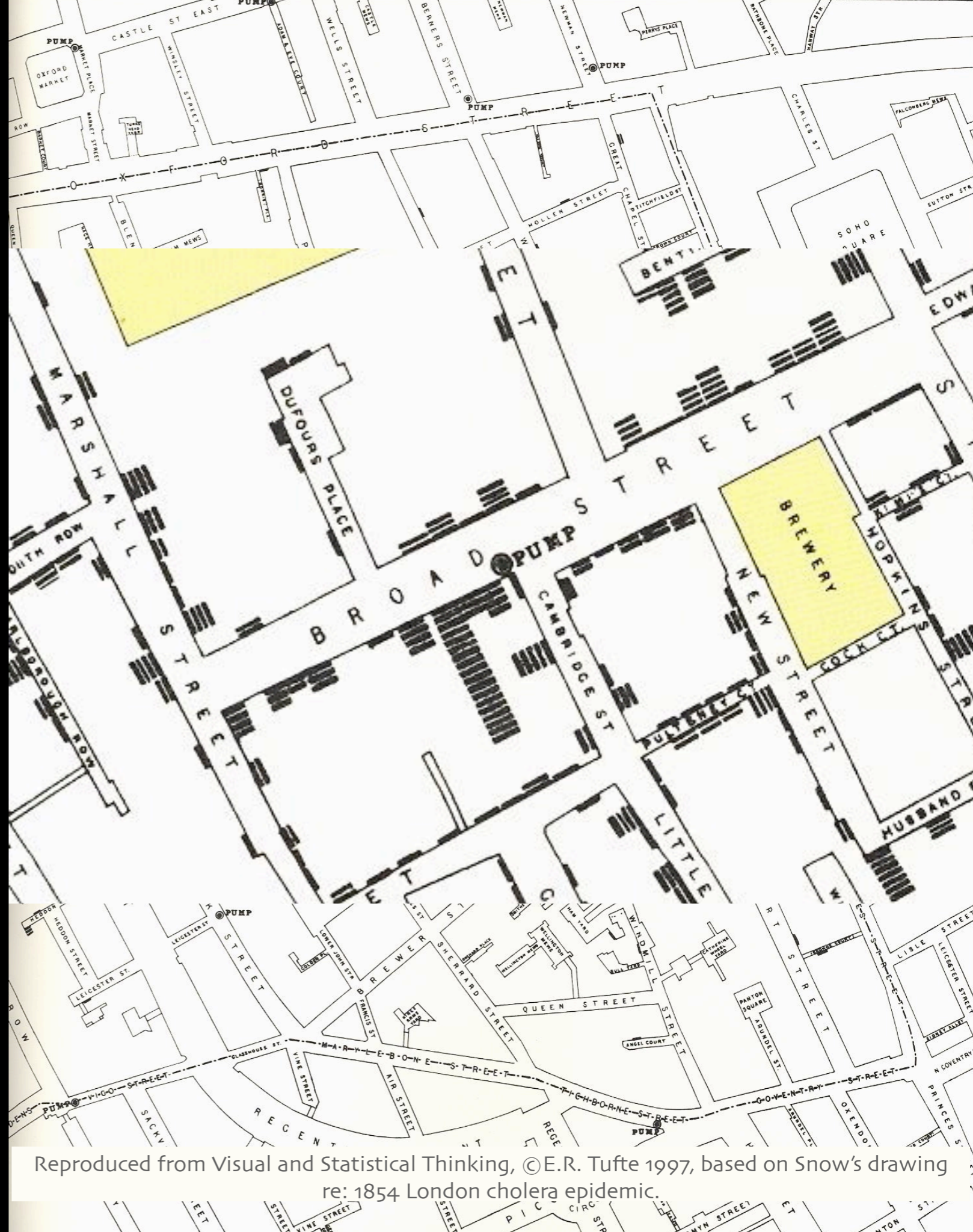
modular craftsmanship, linked views

"Easier"



2011

Craftsmanship (in 1854)



Reproduced from Visual and Statistical Thinking, © E.R. Tufte 1997, based on Snow's drawing re: 1854 London cholera epidemic.

Craftsmanship (in 1854)

Displaying
“high-dimensional” data

with

“multi-functioning
graphical elements”



Reproduced from Visual and Statistical Thinking, © E.R. Tufte 1997, based on Snow's drawing re: 1854 London cholera epidemic.

What Computers Can Let us Craft (2008)

Elements...

✓ Maps

✓ Tables

✗ Graphs

✓ Charts

✓ Illustrations

✓ Combinations

Live Scoreboard | Celtics.com

SCOREBOARD

DEN 116	WAS 72	POR 97	PHI 46	MIL 34	DAL 26-11	LAL 25-11
CHA 119	BOS 79	NJN 70	SAS 52	UTA 34	SAC 14-21	SEA 9-27
FINAL 2:34	4th 0:50	4th	Halftime 5:36	2nd 10:00	10:00	10:00

COURTSIDE LIVE

19-16 STANDINGS

Fouls 1, Full 4, :20

02:46

1 2 3 4 OT T

18	17	24	13		72
18	19	26	16		79

Fouls 1, Full 3, :20

30-5 STANDINGS

COURTSIDE LIVE | BOX SCORE | PLAY-BY-PLAY | Highlights | Watch the Game | Listen to the Game

WAS SELECT: ○ ALL ● ACTIVE 5

PLAYER NAME	PTS	REB	AST	F
Daniels, Antonio	7	2	8	0
Steuenson, DeSha	11	3	4	2
Jamison, Antawn	18	10	0	3
Butler, Caron	14	3	1	3
Haywood, Brenda	12	5	0	3
Blatche, Andray	3	5	0	3
Mason, Roger	3	1	1	5
Songaila, Darius	2	1	1	2
Young, Nick	2	0	0	0
Pecherou, Oleksiy	0	1	0	0
Arenas, Gilbert				
McGuire, Dominic				

BOS SELECT: ○ ALL ● ACTIVE 5

PLAYER NAME	PTS	REB	AST	F
Rondo, Rajon	4	2	2	2
Allen, Ray	16	6	3	2
Garnett, Kevin	21	6	6	3
Pierce, Paul	16	4	2	3
Perkins, Kendrick	9	3	1	3
House, Eddie	5	6	3	1
Allen, Tony	4	4	0	0
Davis, Glen	1	0	0	2
Posey, James	3	2	0	2
Pollard, Scott				
Scalabrino, Brian				
Powe, Leon				

TD Banknorth GARDEN

WIZARDS

CELTICS

WAS show: ● made shots ✓ X missed shots ✓

BOS show: ● made shots ✓ X missed shots ✓



What Computers (D3) Can Let us Craft (2012)

UPDATED June 12, 2012

N.B.A. FINALS

[f RECOMMEND](#) [TWITTER](#) [LINKEDIN](#) [SIGN IN TO E-MAIL](#) [SHARE](#)

Where the Heat and the Thunder Hit Their Shots

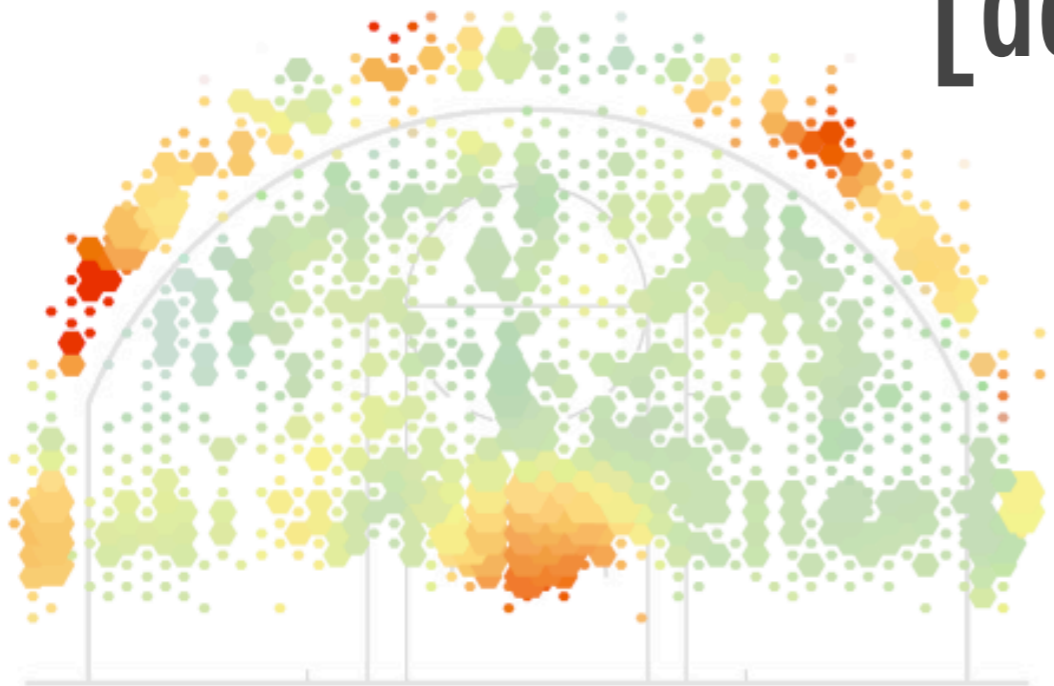
The shooting patterns for the players on the Miami Heat and the Oklahoma City Thunder reveal where they are most dangerous on the court. Below, compare each player's strengths using court maps and analysis by Kirk Goldsberry, a geography professor at Michigan State. [Related Article »](#)

All Shots | **3-Pointers** | **Midrange** | **Close Range** | **Number of attempts** (Low to High) | **Points per region** (Low to High)

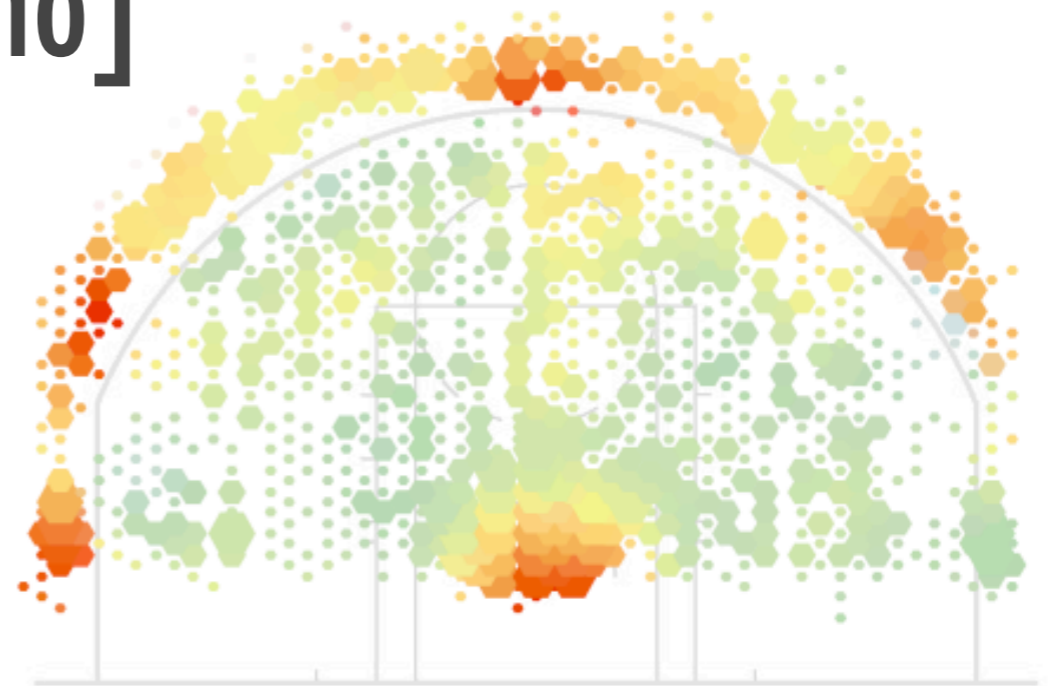
Miami Heat | **Oklahoma City Thunder**

TOTAL SHOTS **5,209** | POINTS PER SHOT **1.01** | F.G. PERCENT **47%** | TOTAL SHOTS **5,228** | POINTS PER SHOT **1.03** | F.G. PERCENT **47.1%**

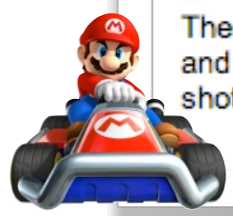
[demo]



The Heat rely on player positioning to create isolation plays for LeBron James and Dwyane Wade, often on the left side. The Heat take many fewer 3-point shots than the Thunder.



The Thunder are effective from almost any area on the court and shoot many more 3-point shots than the league average. Kevin Durant and James Harden are potent from the top of the arc.



Data • Dimensions • Display

What...

...is easier now than before?

fast computation, animation, 3D

...was easier before than now?

craftsmanship

...should be easier in the future?

modular craftsmanship, linked views

What...

...is easier now than before?

fast computation, animation, 3D

...was easier before than now?

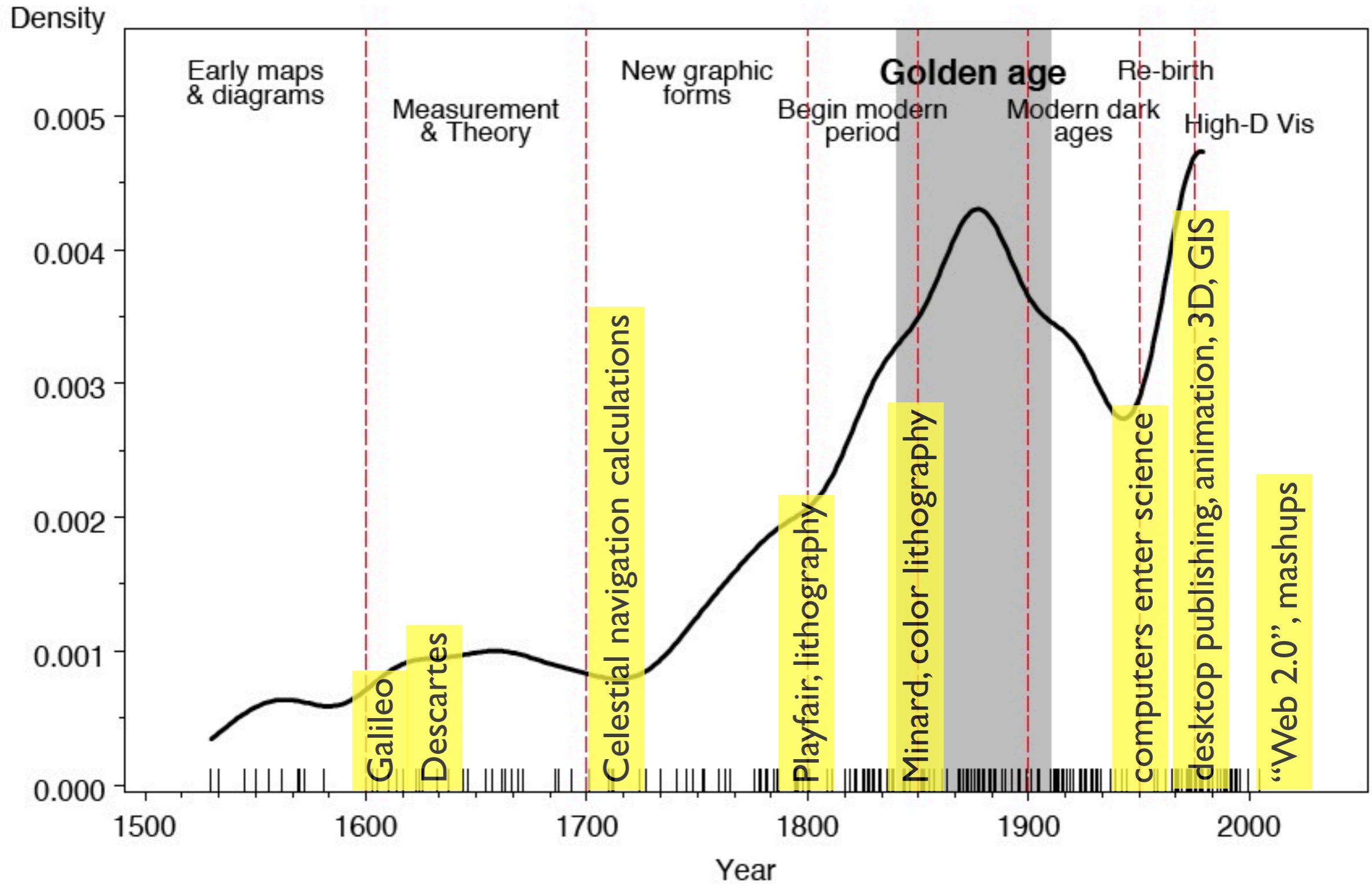
craftsmanship

...should be easier in the future?

modular craftsmanship, linked views

First, the past...

Milestones: Time course of developments



adapted from Friendly, "The Golden Age of Statistical Graphics," *Statistical Science*, 2009

Galileo Galilei (1564-1642)

Sc. Principale.

Galileo Galilei, Familiare Seruo della Ser. V. inuigilante
 do assiduo, et lo ogni spirito se bene no solo satisfar
 aluano che non della lettera di Madonata nelle sue
 Dio di Padova,

Inuere d'auere determinato di presentare al Sc. Principe
 l'occhio et il p. essere di giuamenti inestimabile di ogni
 negozio et in circa marittima o terrestre stimo di tenere que
 sto nuovo artificio nel maggior segreto et solam a disposizione
 di V. Ser. L. Galileo conato dalle piu uide speculazioni di
 prospettina in l'uantaggio di scoprire Legni et Vele dell' inimico
 di due hore et piu di tempo prima di esse sopra noi et distinguend
 il numero et la qualita de i Vasselli giudicare le sue forze
 ballottarsi alla caccia et combattimento o alla fuga, o pure esser
 nella campagna aperta uedere et particolarmente distinguere ogni suo
 posto et provvedimento.

Adi 7. di gennaio
 Giove si uede a 7. * uici: 10. 11.
 Adi 8. uici * * * * *
 4. * * * * * ora d'uy diretto et no retrogrado
 Adi 12. si uede in tale uisione * * * * *
 N. 13. si uede in uisione a Giove 4 stelle * * * * *
 Adi 14. è angelo * * * * *
 N. 15. * * * * * la pressi a 4. ora in uici la 4. ora di =
 sbante dalla 3. a coppia terra * * * * *
 Lo spazio delle 3. uide uide no om
 maggiore del diametro di 7. et e
 uici in linea retta.

7	* * ○ *	17	* ○
8	○ * * *	18	* ○
10	* * ○	19	* ○ * *
11	* * ○	19	* ○ * *
12	* ○ *	20	○ * ○ ○
13	* ○ * *	21	... ○ *
15	○ * * * *	22	* ○ * *
15	○ * * *	22	* ○ * *
16	* ○ *	23	* ○ *
17	* ○ *	24	* ○ *
		24	* ○ *

SIDERIUS NUNCIUS

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

East * ○ * * West

0 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 wa
 distant 3 minutes from the next one, while this one was 40 second
 from Jupiter; Jupiter was 4 minutes from the nearest western one
 and this one 6 minutes from the westernmost one. Their magnitude
 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 ** ○ * * 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.

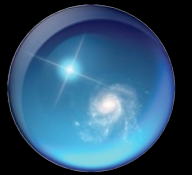
On the fifth, the sky was cloudy.

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

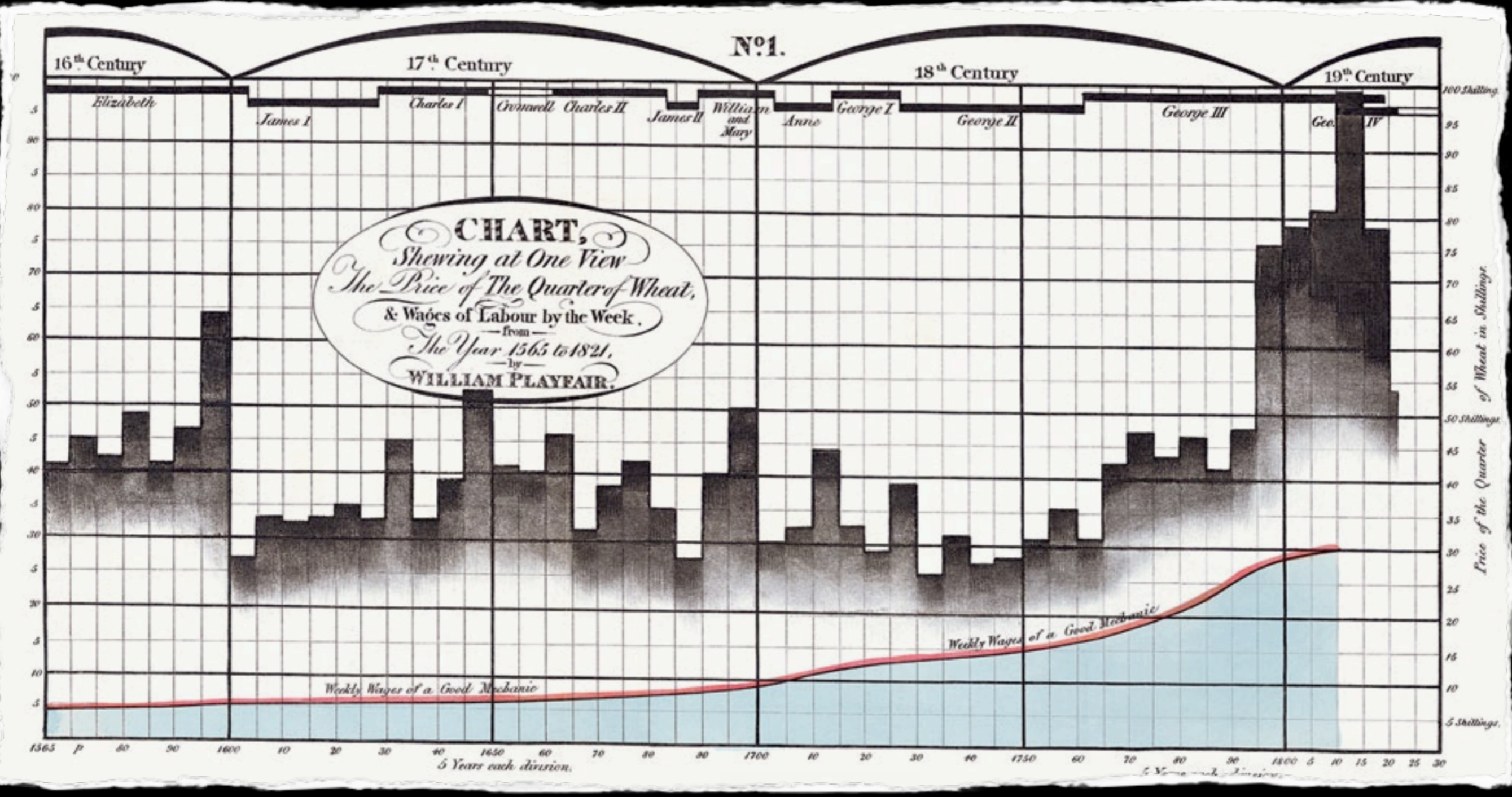
East * ○ * West

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.

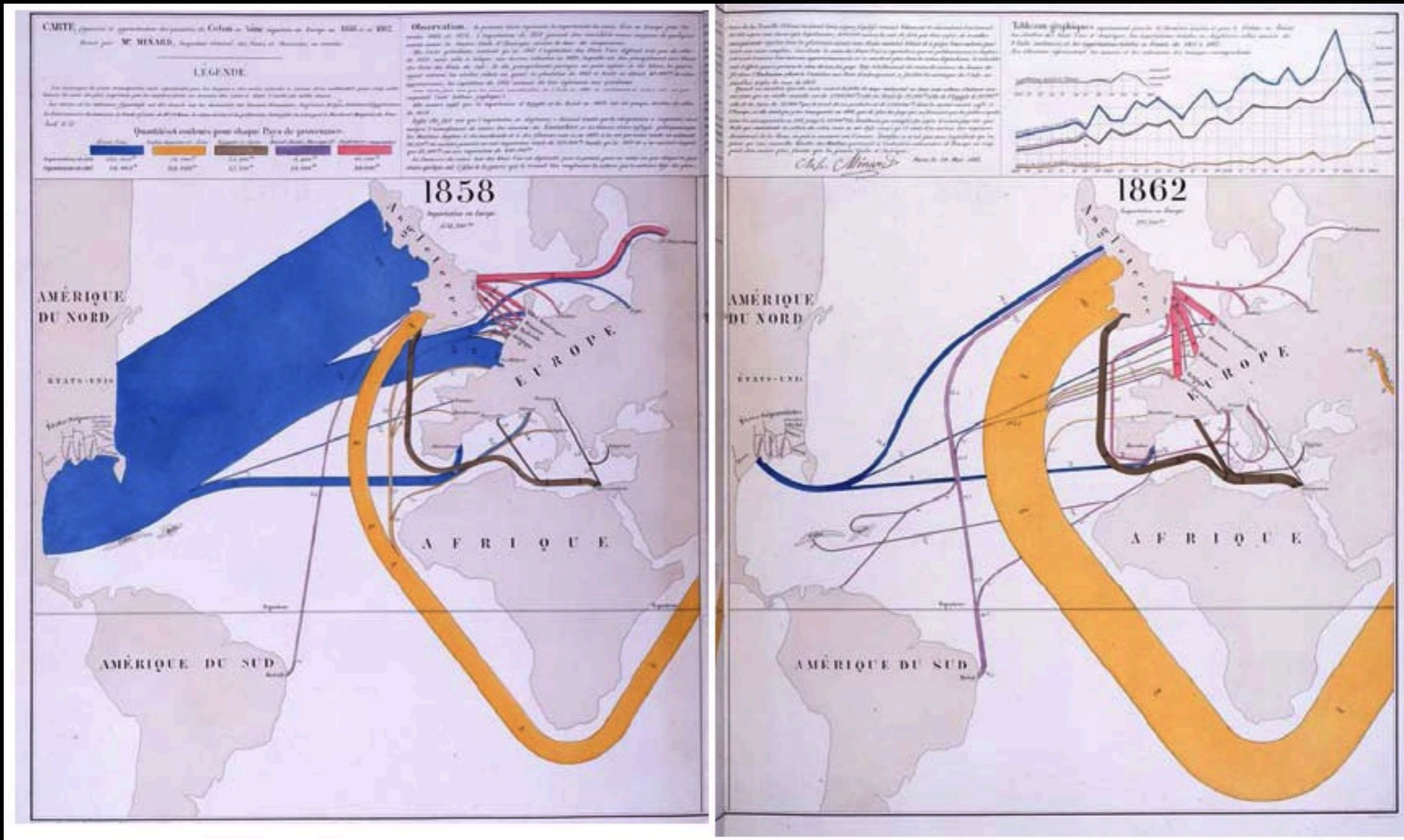
On the seventh, two stars stood near Jupiter both to the east



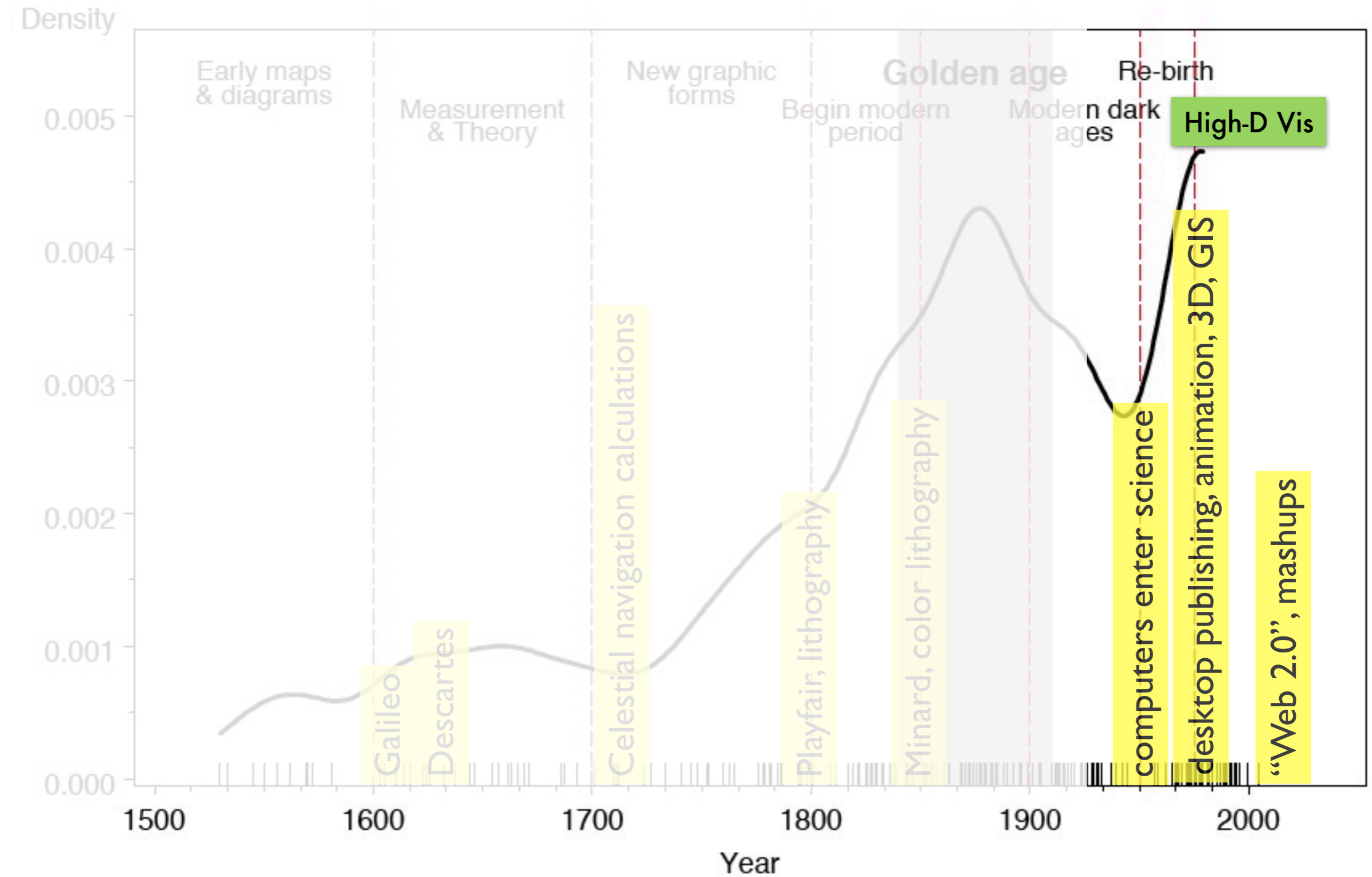
William Playfair (1759-1823)



Charles Joseph Minard, in color (1781-1870)



Milestones: Time course of developments

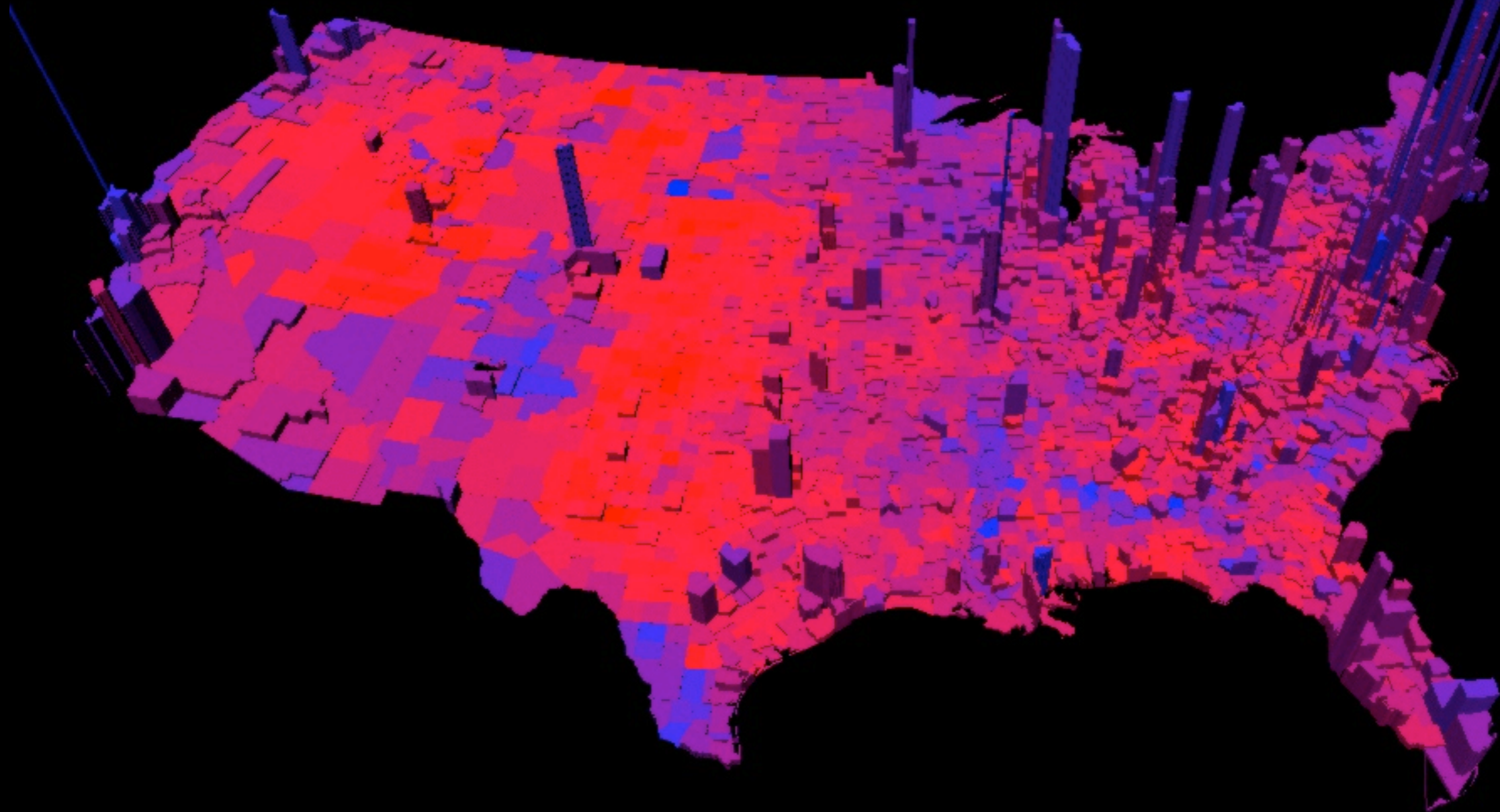


adapted from Friendly, "The Golden Age of Statistical Graphics," *Statistical Science*, in press (2008)

High-D Vis

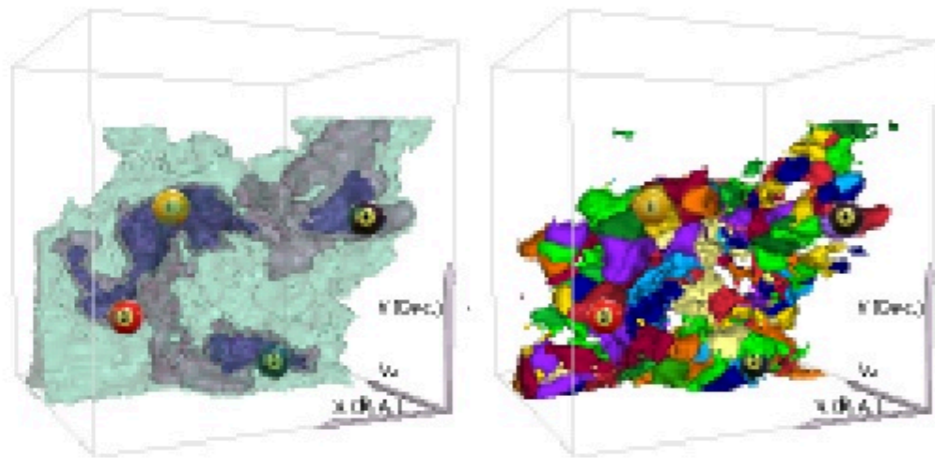
Data • Dimensions • Display

“High-dimensional” or “Multivariate” Data and High(er) Dimensional Displays

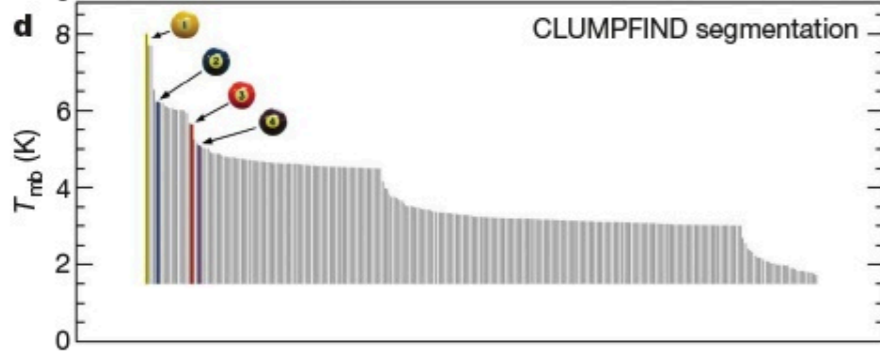
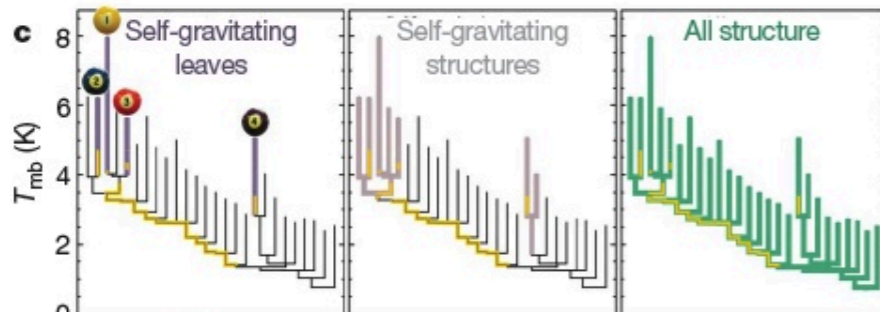


"High-dimensional" or "Multivariate" Data (Astronomy=Biology)

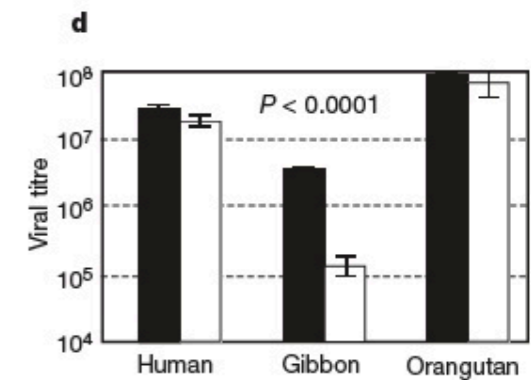
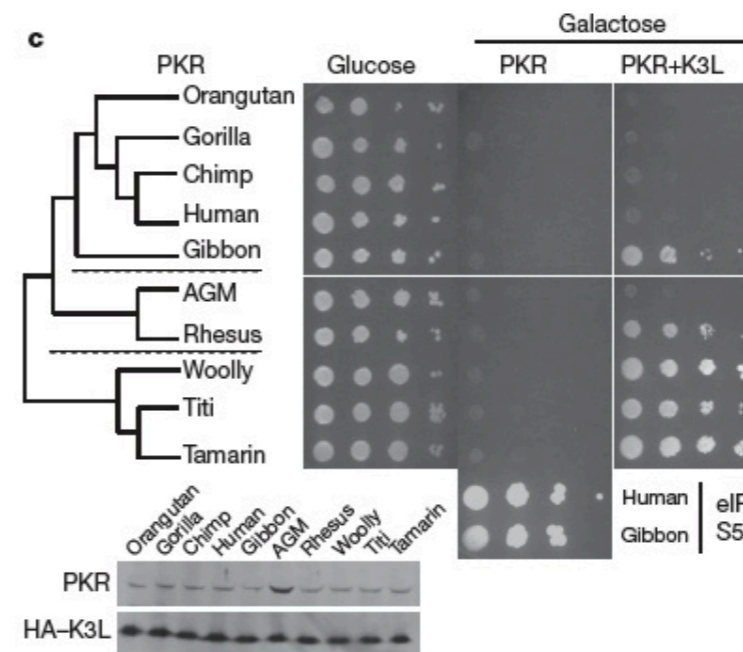
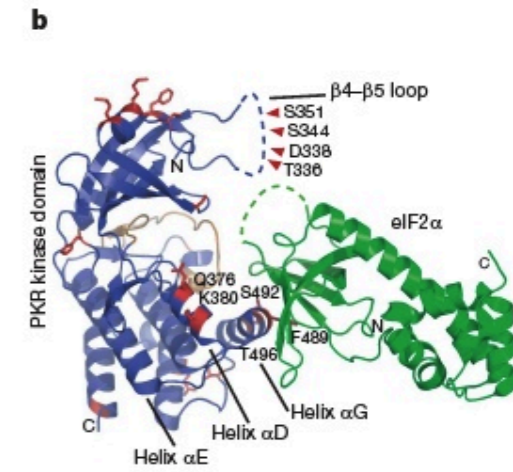
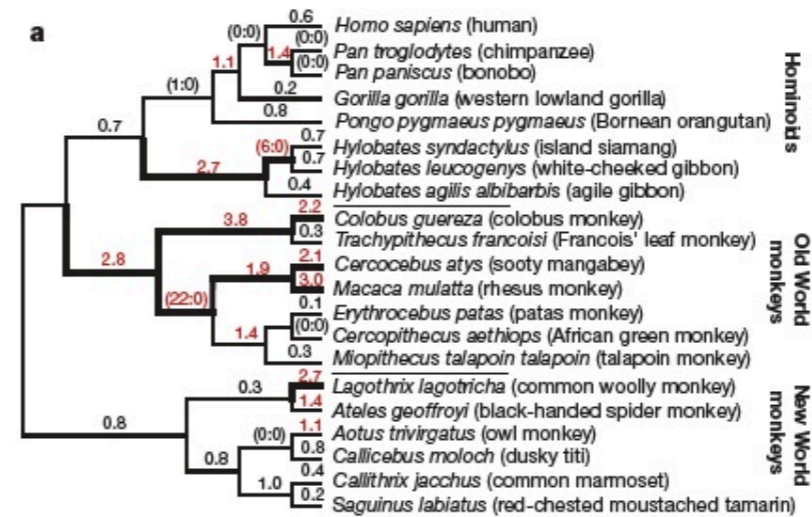
LETTERS



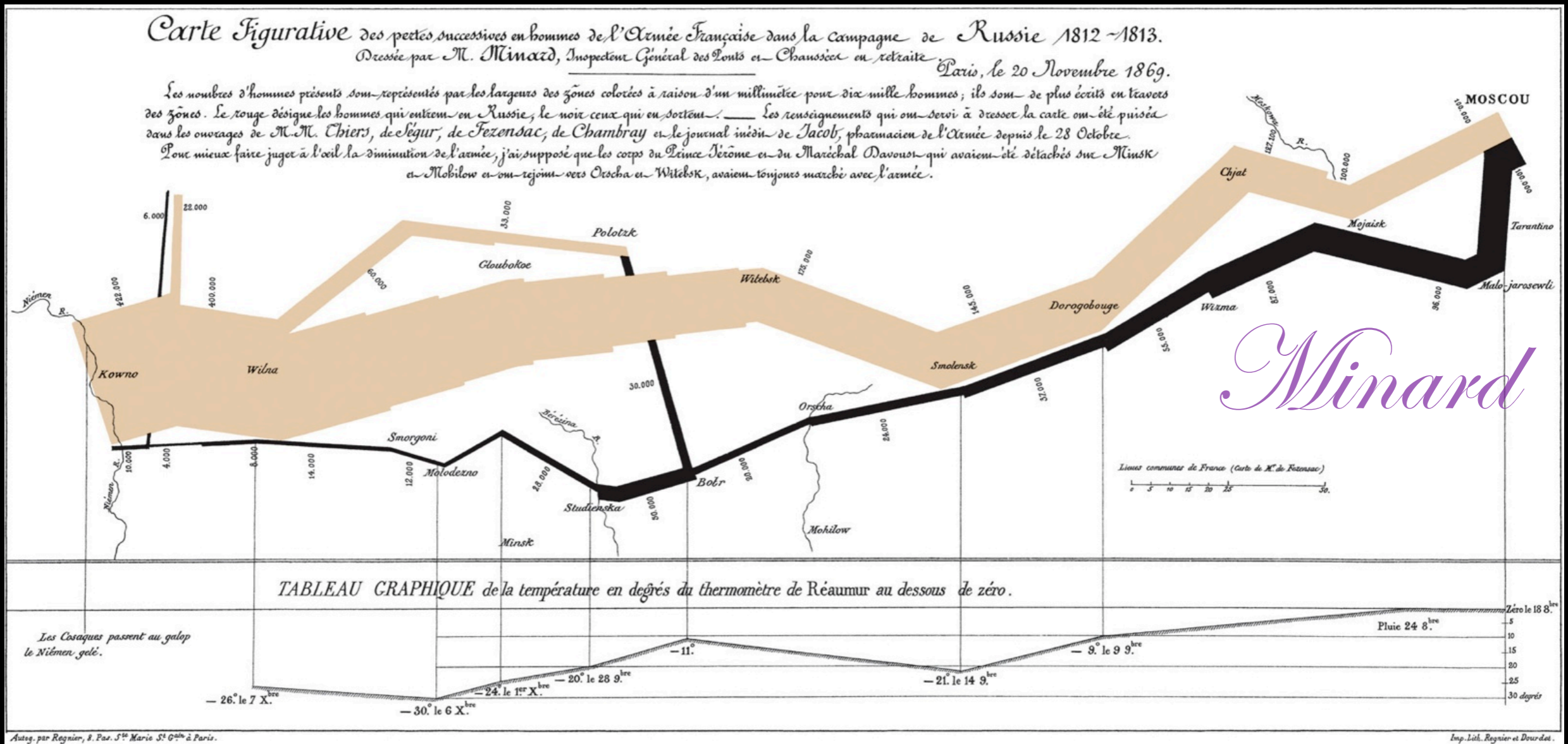
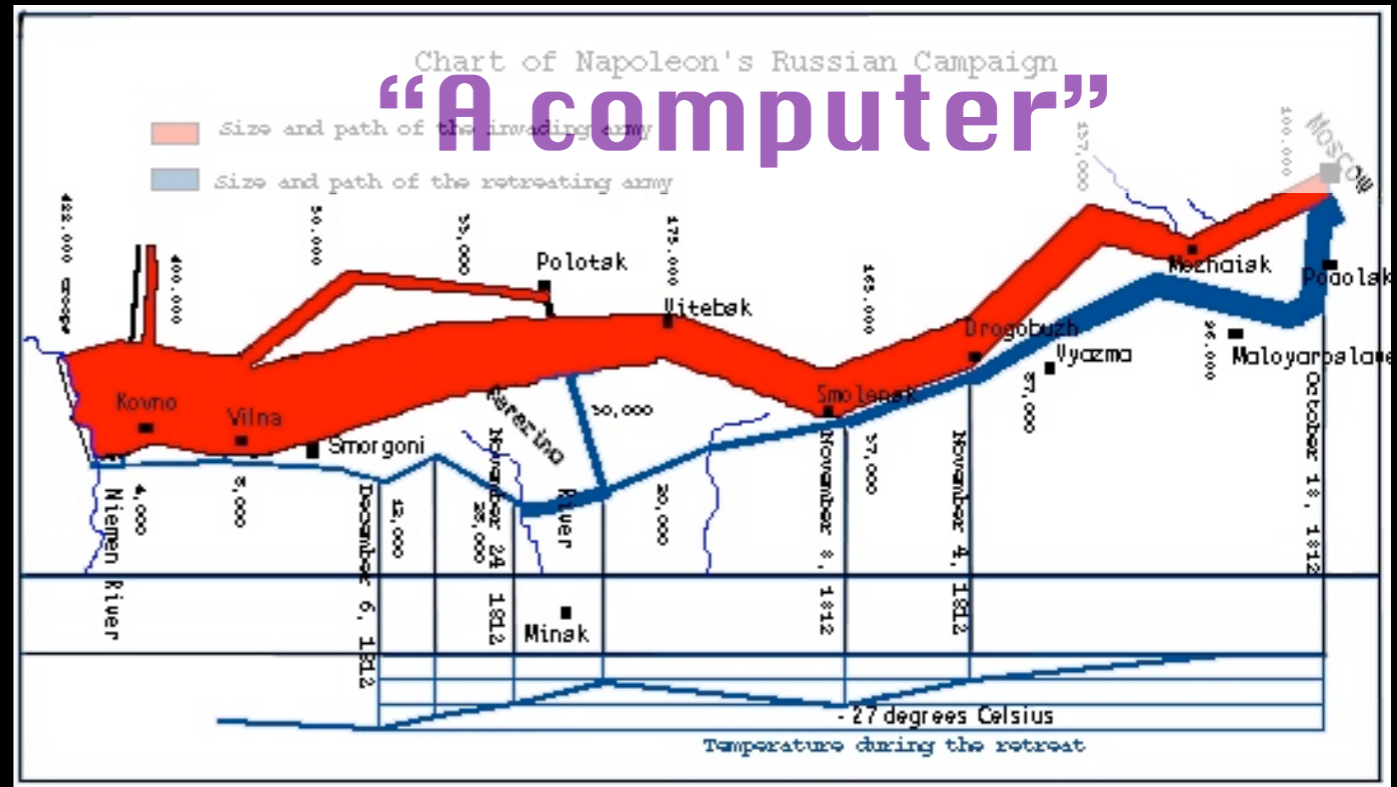
Click to rotate



LETTERS



How much are we held back today by digital tools?



How to advance the digital (visualization) tools for quantitative research?

SEAMLESS ASTRONOMY
Linking scientific data, publications, and communities

ABOUT PROJECTS PEOPLE RESOURCES DATAVERSE

SEAMLESS ASTRONOMY

About

The Seamless Astronomy Group at the Harvard-Smithsonian Center for Astrophysics brings together astronomers, computer scientists, information scientists, librarians and visualization experts involved in the development of tools and systems to study and enable the next generation of online astronomical research.

Current projects include research on the development of systems that seamlessly integrate scientific data and literature, the semantic interlinking and annotation of scientific resources, the study of the impact of social media and networking sites on scientific dissemination, and the analysis and visualization of astronomical research communities. Visit our project page to find out more.

Viz-e-Lab

Projects 2011

Taste-Testing

WorldWide Telescope/Ambassadors

Linked Views/Dendrograms

Viz

Seamless Astronomy

Wolbach User Experience Lab

The Dataverse Network Project

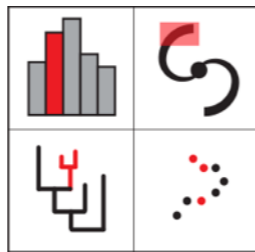
VAO

glue
multidimensional data exploration

collaborators/contacts at CfA

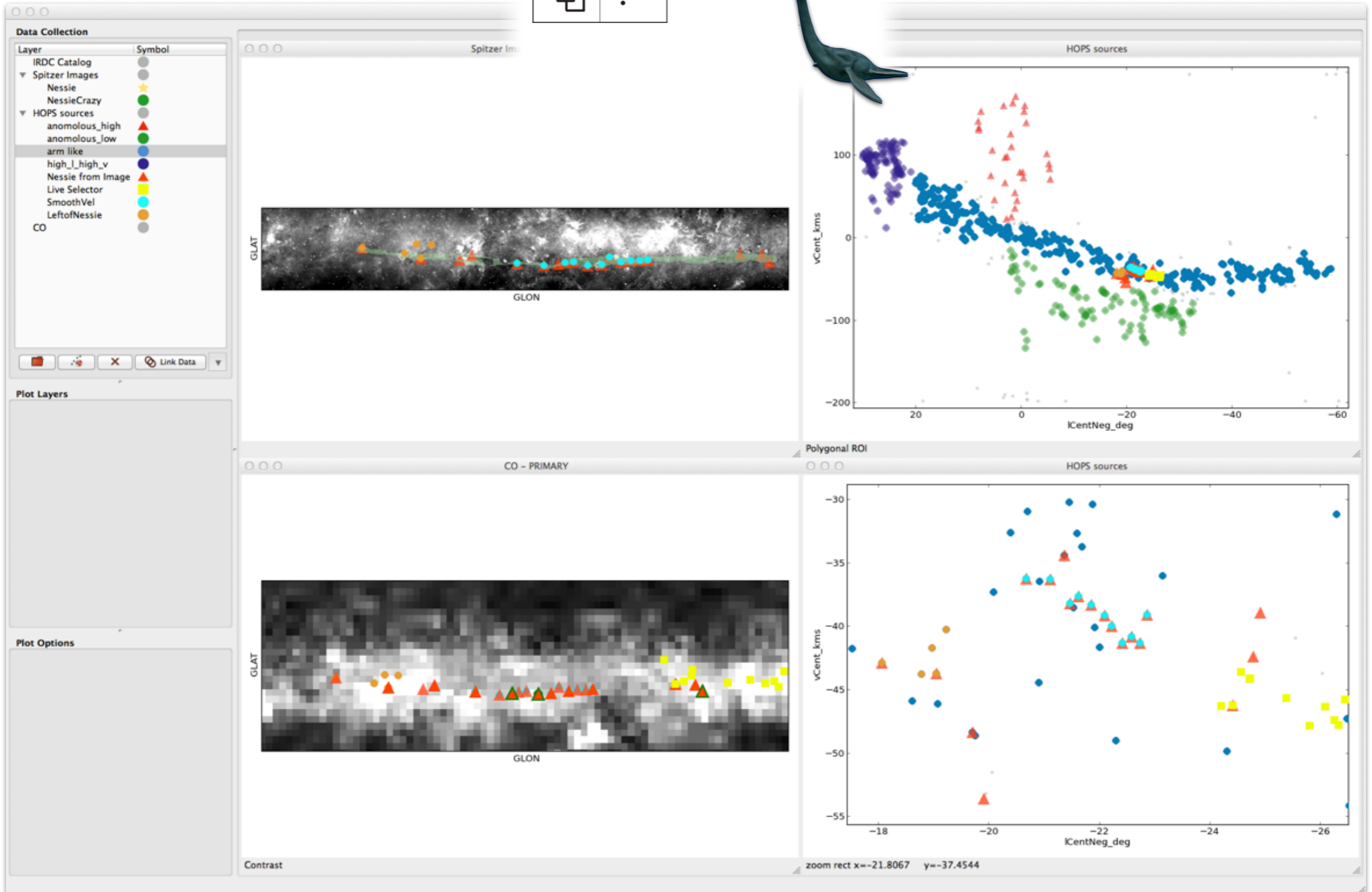
Seamless Astronomy: Alyssa Goodman Online Astronomy Group, CfA Data Archives: Gus Muench ADS Group: Alberto Accomazzi
WorldWide Telescope Ambassadors: Pat Udomprasert High-Dimensional Data Visualization & Interaction: Michelle Borkin
Wolbach Library Lab at CfA: Christopher Erdmann VAO at CfA: Pepi Fabbiano Social Networks in Science: Alberto Pepe
Questions about using the Viz-e-Lab? Contact Sarah Block, 5-7331, sblock@cfa.harvard.edu





glue

multidimensional data exploration



Glue collaboration: **Beaumont**, Borkin, Goodman, Pfister, Robitaille

The AstroMed (Back) Story



Themes	TED Conferences	TED
Speakers	TEDx Events	TED
Talks	TED Prize	
Translations	TED Fellows	

TED Fellows The TED Fellows Directory > Michelle Borkin 2009



Michelle Borkin is now a SEAS PhD Student, advised by Profs. Alyssa Goodman (Astronomy) and Hanspeter Pfister (SEAS), and IIC +AstroMed became the bases for the Viz-e-Lab



2011 Visual Business Intelligence

A blog by Stephen Few

Home About Consulting Workshops Courses Examples Library **Blog** Discuss

VisWeek 2011 – Award-Worthy Visualization Research

On Tuesday in this blog I expressed my frustration with VisWeek's information visualization research awards process. I don't want to leave you with the impression, however, that the state of information visualization research is bleak. Each year at VisWeek I find a few gems produced by thoughtful, well-trained information visualization researchers. They identified potentially worthy pursuits and did well-designed research that produced useful results. While puzzling over the criteria that the judges must have used when selecting this year's best paper, I spent a few minutes considering the criteria that I would use were I a judge, and came up with the following list with points totaling to 100:

Effectiveness (It does what it's supposed to do and does it well.) — 30 points

Usefulness (What it does addresses real needs in the world.) — 30 points

10 points

ses.) — 10 points

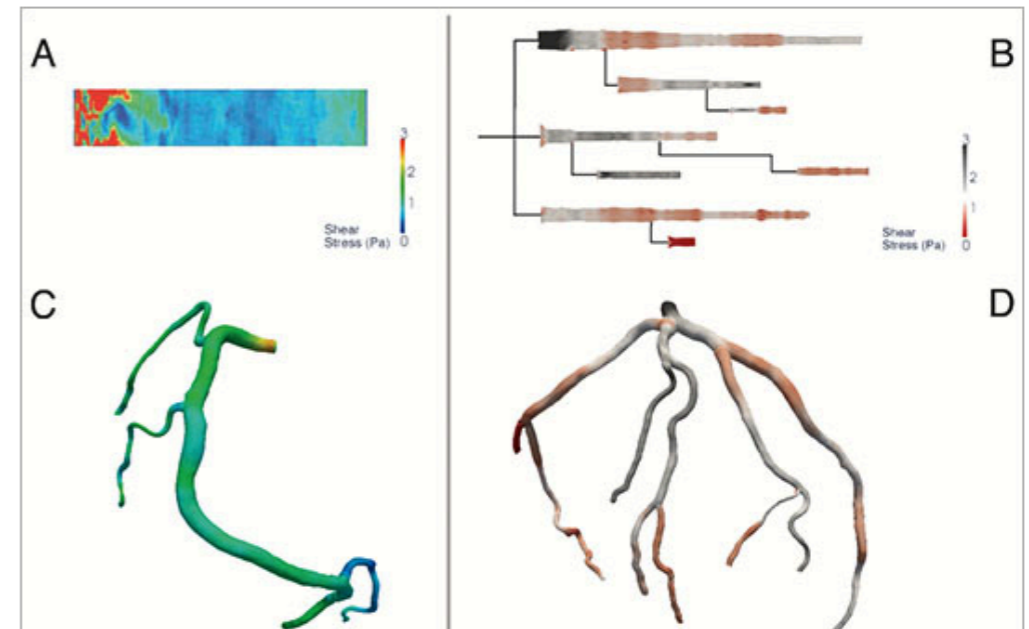
ew way.) — 10 points

e.) — 10 points

to some degree, but this gives you an idea of the importance of each.

e by its elegance and exceptional usefulness Harvard University's School of Engineering and

Applied Sciences titled "Evaluations of Artery Visualizations for Heart Disease Diagnosis."



TEDGlobal 2009

AstroMed09

The Inaugural Sydney International Workshop on Synergies in Astronomy and Medicine

14–16 December, 2009
The University of Sydney

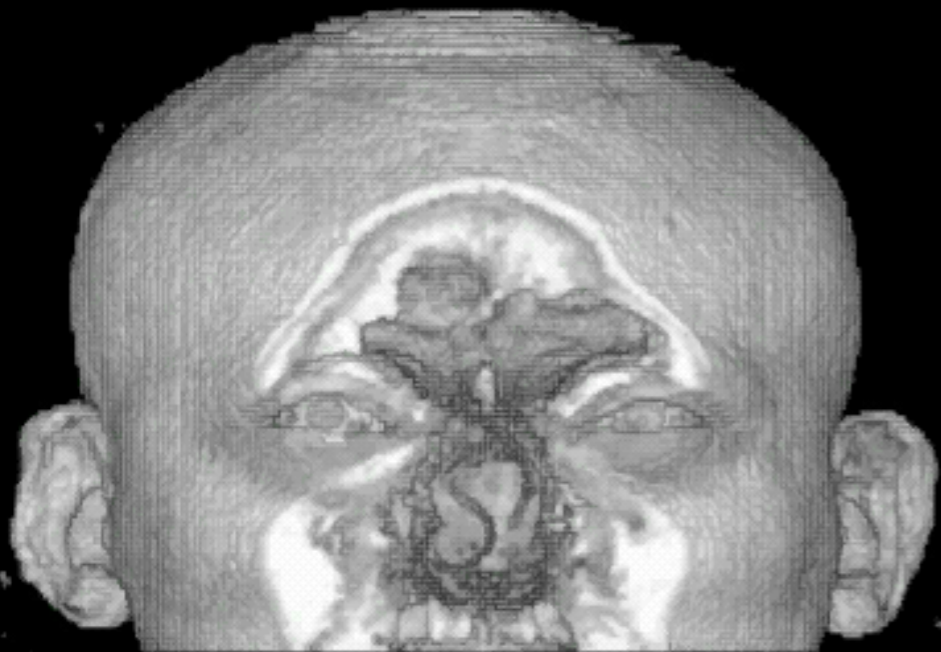
Bio

Michelle Borkin interdisciplinary and image analysis. She wrote her work on the application of astronomical data as part of the "AstroMed" project at Harvard's Initiative for Data-Driven Works with the development of tools to improve their effectiveness in multiple

serting a stent

"Astronomical Medicine"

"KEITH"



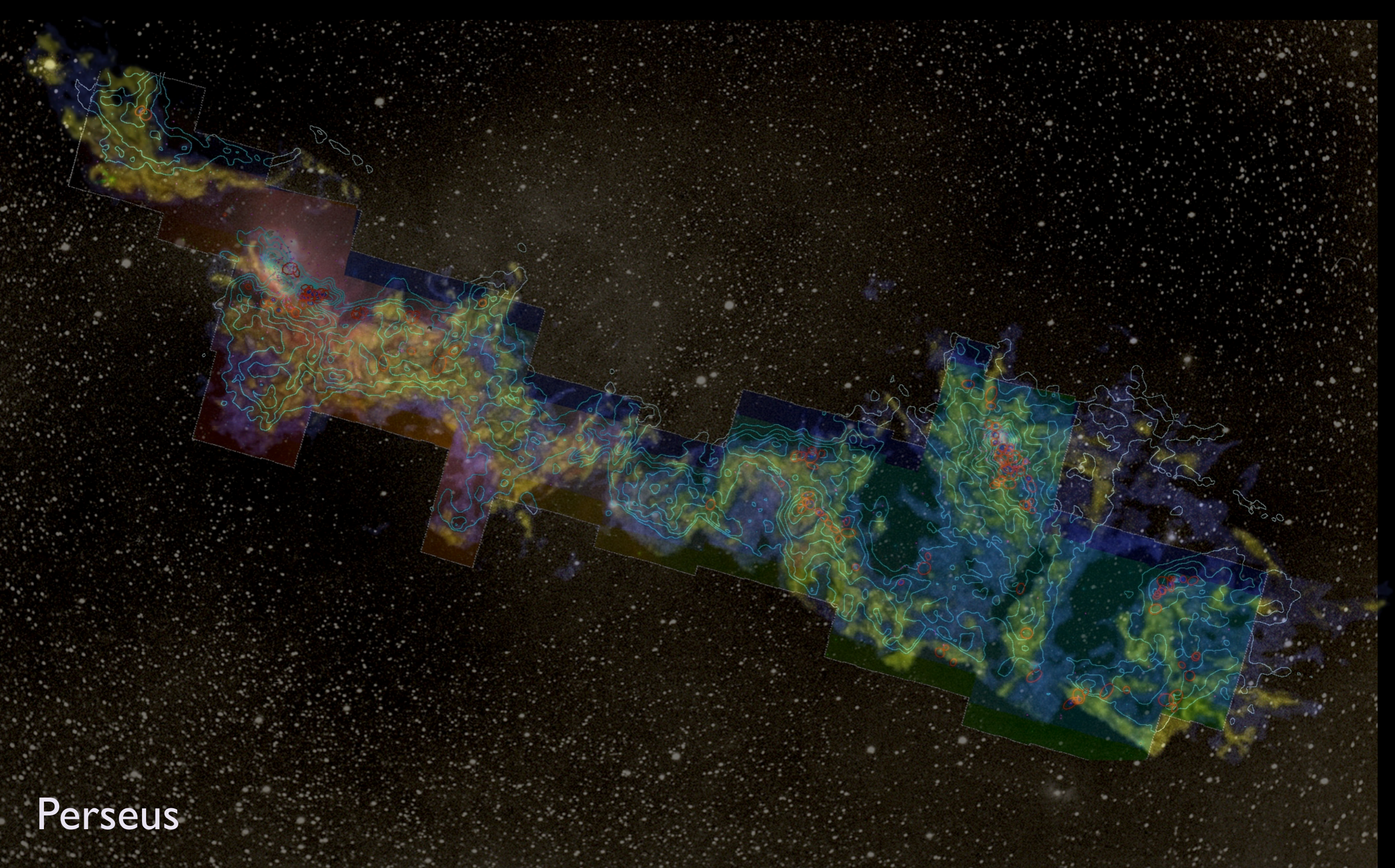
"z" is depth into head

"PERSEUS"



"z" is line-of-sight velocity

(This kind of "series of 2D slices view" is known in the Viz as "the grand tour")



Perseus

3D Viz made with VolView

How interactive? How "linked"?

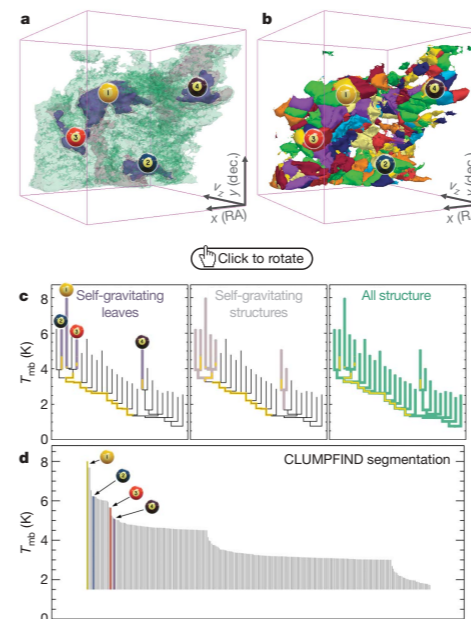


Figure 2 | Comparison of the 'dendrogram' and 'CLUMPFIND' feature-identification algorithms as applied to ^{13}CO emission from the L1448 region of Perseus. **a**, 3D visualization of the surfaces indicated by colours in the dendrogram shown in **c**. Purple illustrates the smallest scale self-gravitating 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 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 set⁸ 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 this early 2D work as inspiration, we have developed a structure-identification algorithm that abstracts the hierarchical structure of a 3D (p - p - v) data cube into an easily visualized representation called a 'dendrogram'¹⁰. Although well developed in other data-intensive fields^{11,12}, it is curious that the application of tree methodologies so far in astrophysics has been rare, and almost exclusively within the area of galaxy evolution, where 'merger trees' are being used with increasing frequency¹³.

Figure 3 and its legend explain the construction of dendrograms schematically. The dendrogram quantifies how and where local maxima of emission merge with each other, and its implementation is explained in Supplementary Methods. Critically, the dendrogram is determined almost entirely by the data itself, and it has negligible sensitivity to algorithm parameters. To make graphical presentation possible on paper and 2D screens, we 'flatten' the dendrograms of 3D data (see Fig. 3 and its legend), by sorting their 'branches' to not cross, which eliminates dimensional information on the x axis while preserving all information about connectivity and hierarchy. Numbered 'billiard ball' labels in the figures let the reader match features between a 2D map (Fig. 1), an interactive 3D map (Fig. 2a online) and a sorted dendrogram (Fig. 2c).

A dendrogram of a spectral-line data cube allows for the estimation of key physical properties associated with volumes bounded by isosurfaces, such as radius (R), velocity dispersion (σ_v) and luminosity (L). The volumes can have any shape, and in other work¹⁴ we focus on the significance of the especially elongated features seen in L1448 (Fig. 2a). The luminosity is an approximate proxy for mass, such that $M_{\text{lum}} = X_{13\text{CO}} L_{13\text{CO}}$, 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 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 fields⁶, its measured value should only be used as a guide to the longevity (boundedness) of any particular feature.

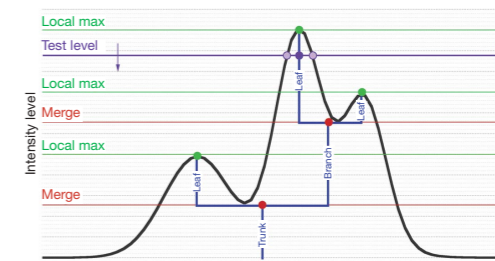


Figure 3 | Schematic illustration of the dendrogram process. Shown is the construction of a dendrogram from a hypothetical one-dimensional emission profile (black). The dendrogram (blue) can be constructed by 'dropping' a test constant emission level (purple) from above in tiny steps (exaggerated in 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, Rosolowsky, Borkin, Foster, Halle, Kauffmann & Pineda, **Nature**, 2009

"Linked Views"

*Contextual,
High-Dimensional
View*

Link

*Flat,
Text-Based
View*

John Tukey's "Four Essentials" (c.1972)

Picturing

Rotation

Isolation

Masking

Selection

and these *"need to work together"*
in a *"dynamic display"*

Brushing

Linking

Results...

1. for immediate **insight**
2. as visual source of **ideas** for statistical algorithms (...relation to SVM)

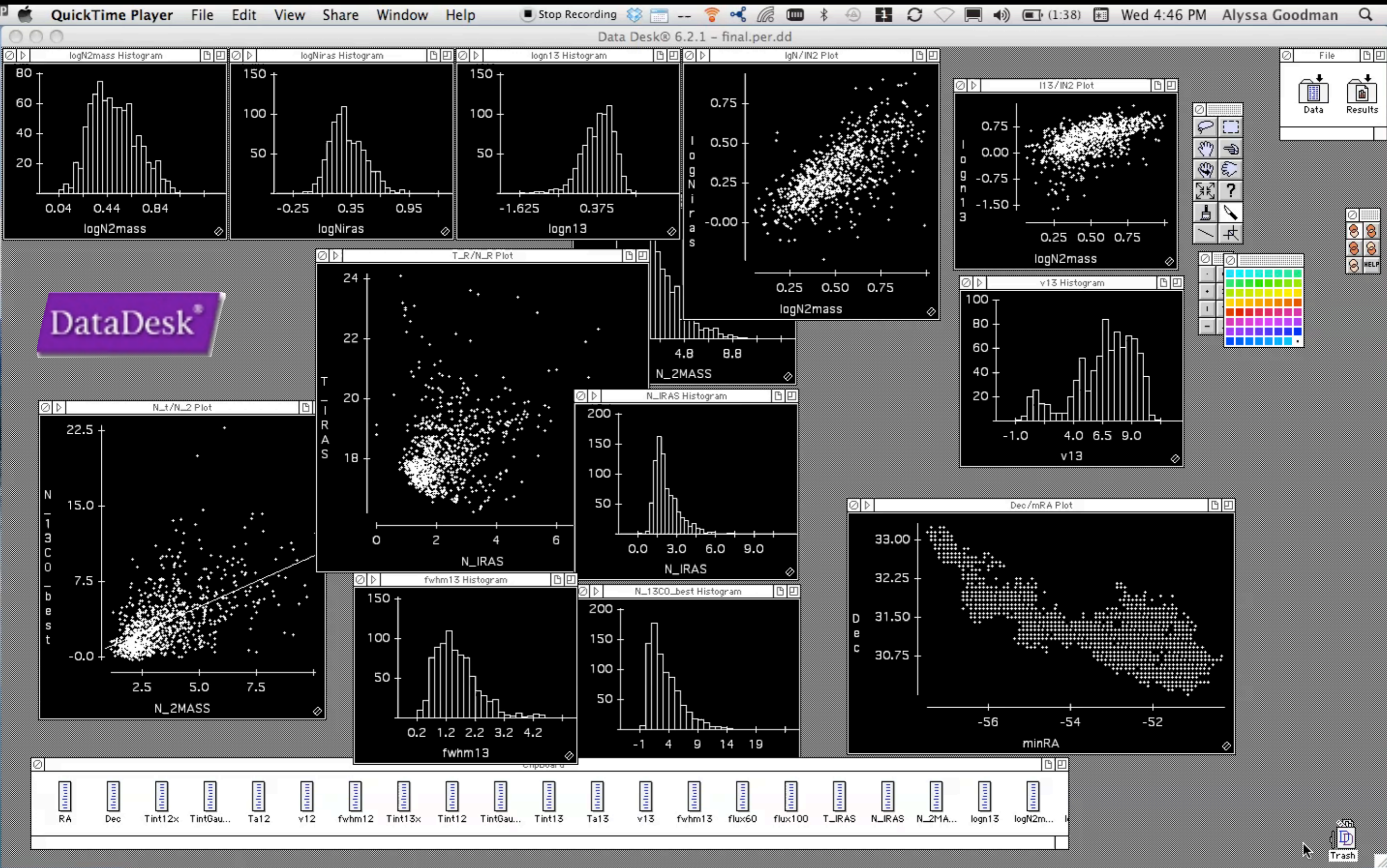
Warning

"details of control can make or break such a system"

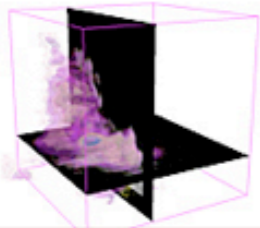
Watch the PRIM-9 video at: <http://stat-graphics.org/movies/prim9.html>



DataDesk (est. 1986)



2008: Dendrostar by Douglas Alan



The Astronomical Medicine Project



Initiative in Innovative Computing at Harvard

Harvard IIC Home

AM Project

- overview
- what's new?
- press
- about us
- contact us

Research

- background
- projects
- papers
- images
- movies

Software

- overview
- Slicer: getting started
- Slicer 3
- fits2itk
- OsiriX
- DendroStar

Links

- Center for Astrophysics
- COMPLETE Survey
- Surgical Planning Lab
- 3D Slicer
- related projects

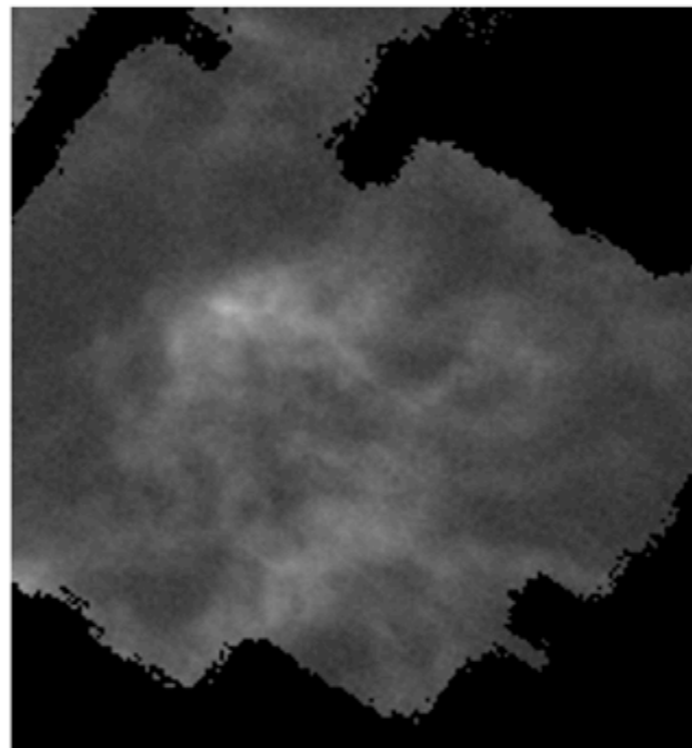
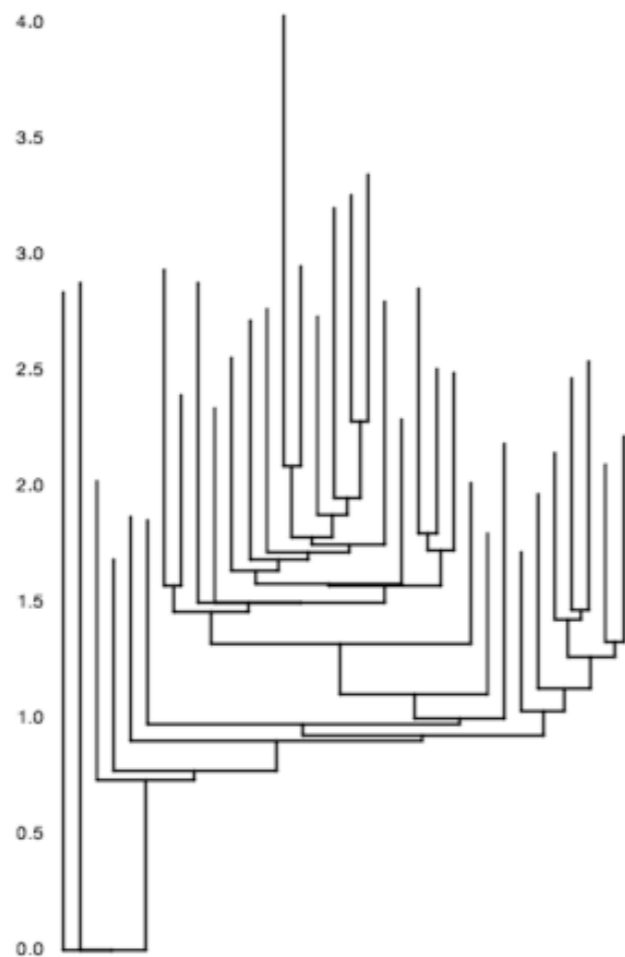
User

- Login

Search

- Titles
- Text

The DendroStar Applet for L1448: Try me!



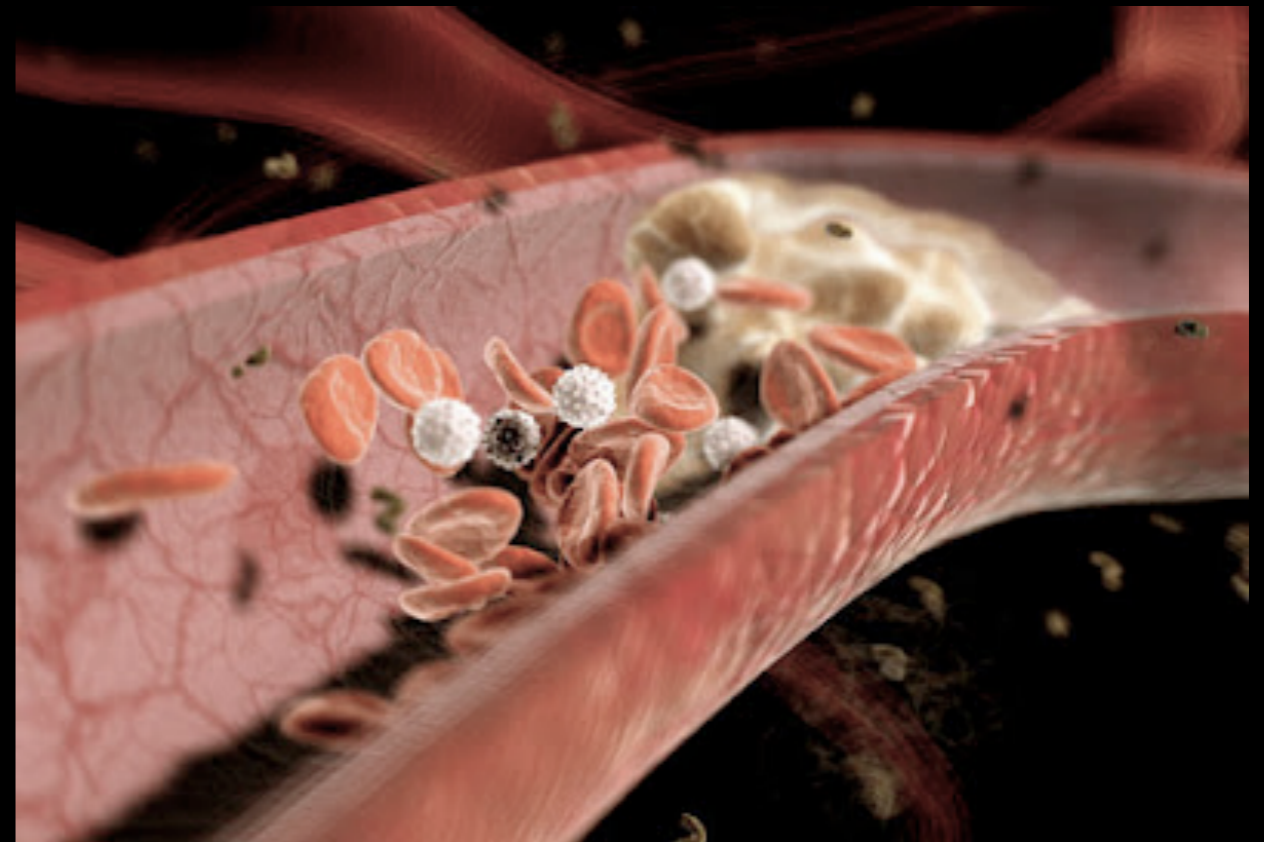
- Tint:
- Suppress tint:
- Reset:

Note: You need to have [Java](#) installed for the applet to work. If the applet doesn't work, try upgrading to a newer version of Java or using [Firefox](#) as your browser.

Click [here](#) for help on using this applet.



2011: The (Medical) Value of Linked Views...



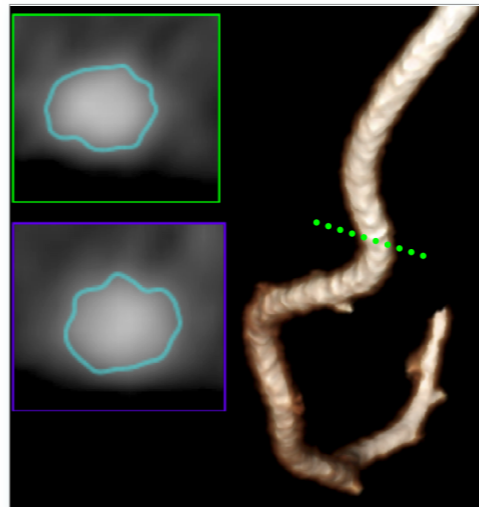
Michelle Borkin

*Harvard School of Engineering & Applied Science Ph.D. student,
supervised by Alyssa Goodman (Astronomer) & Hanspeter Pfister (Computer Scientist)*

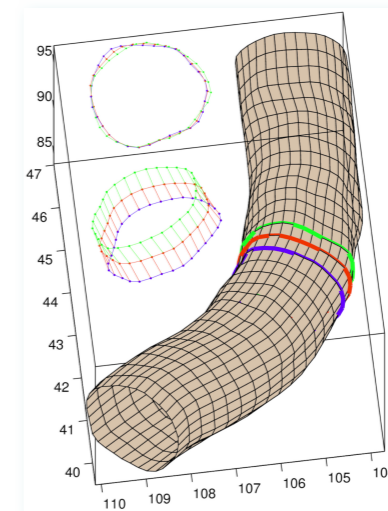
Patients Troubled Hearts, in 3D



Obtain patient CT data



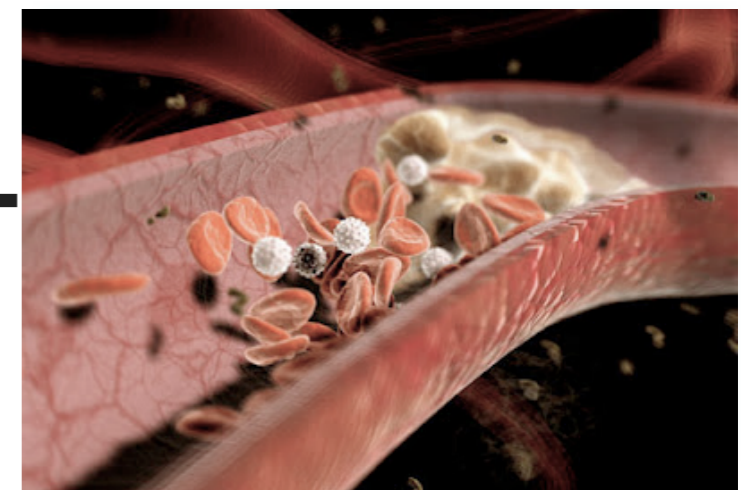
Segment arteries



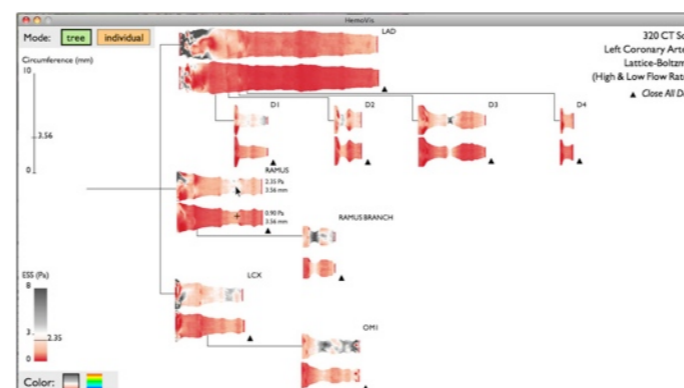
Generate patient geometries



Patient specific flow simulation



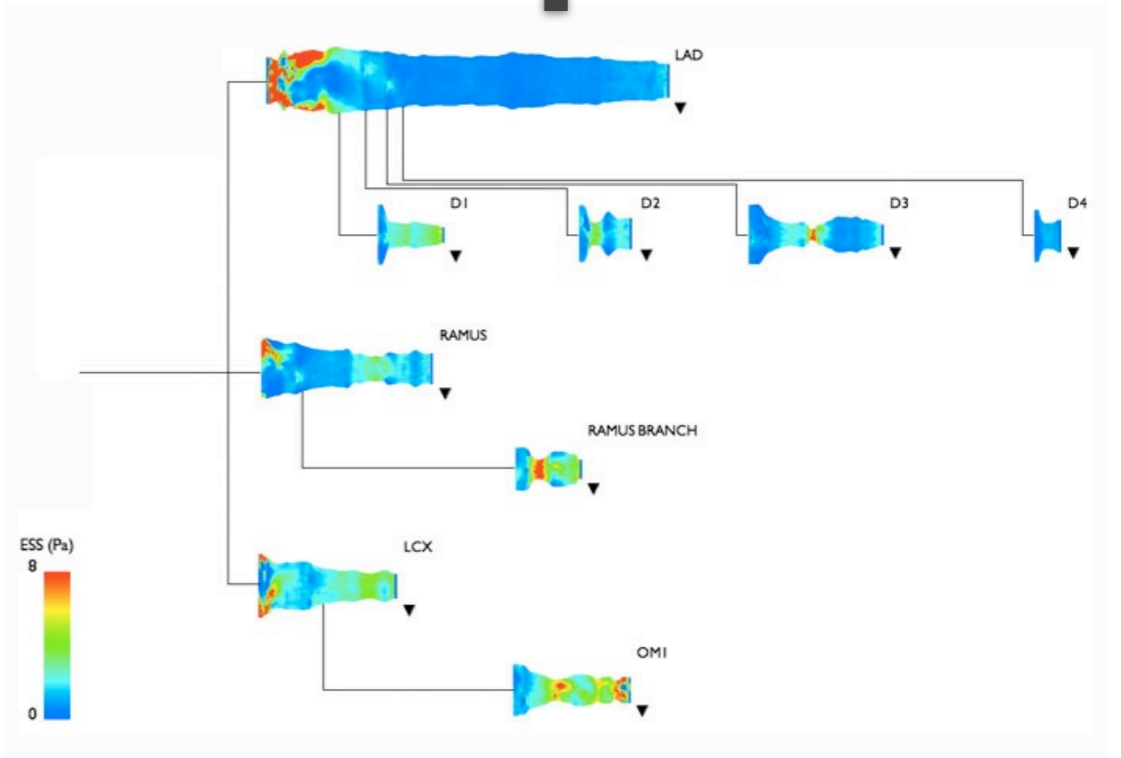
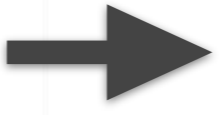
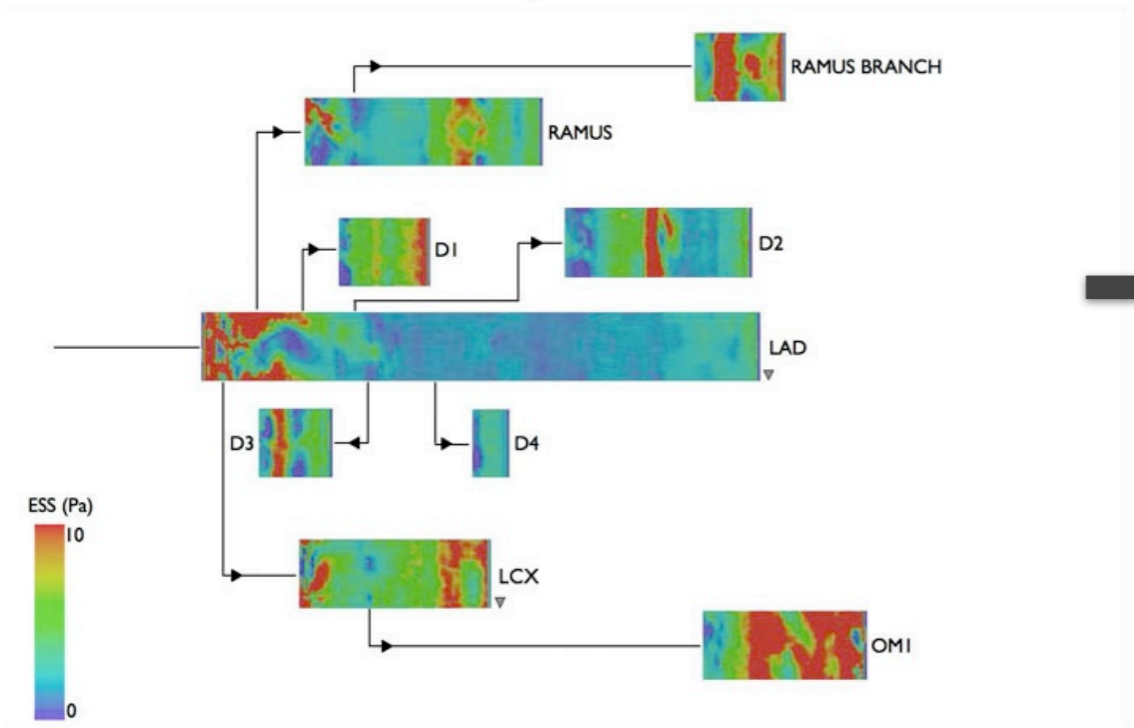
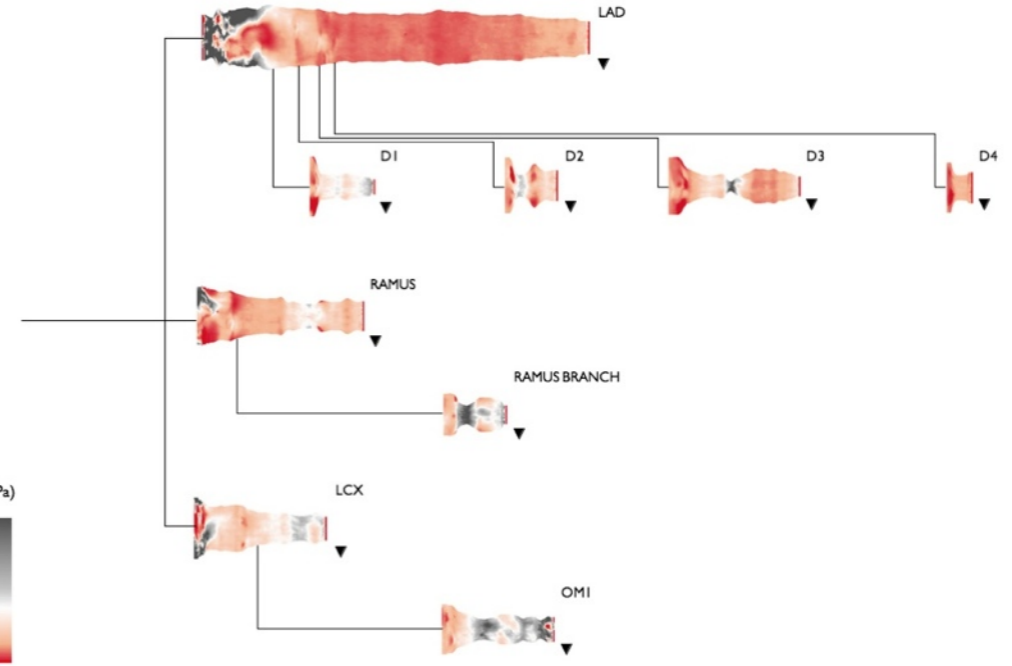
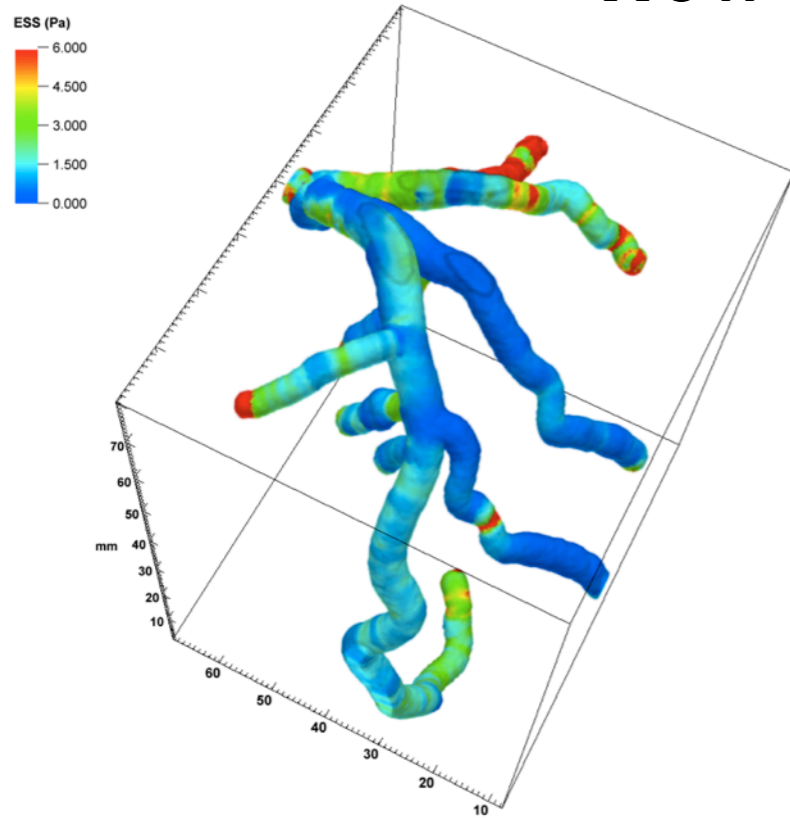
Visualize/analyze data



Clinical decision



How much does viz matter?



ACCURACY

Strong effect of **dimensionality** on accuracy

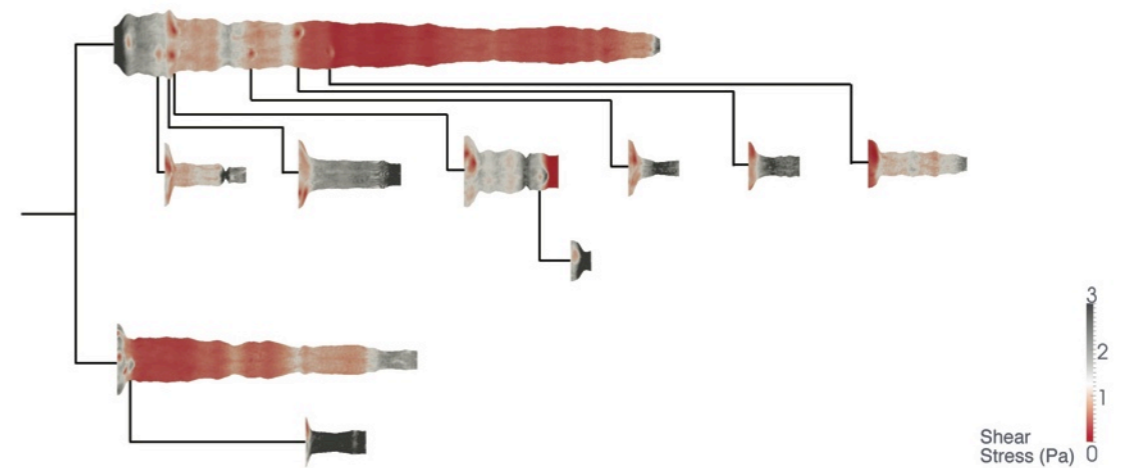
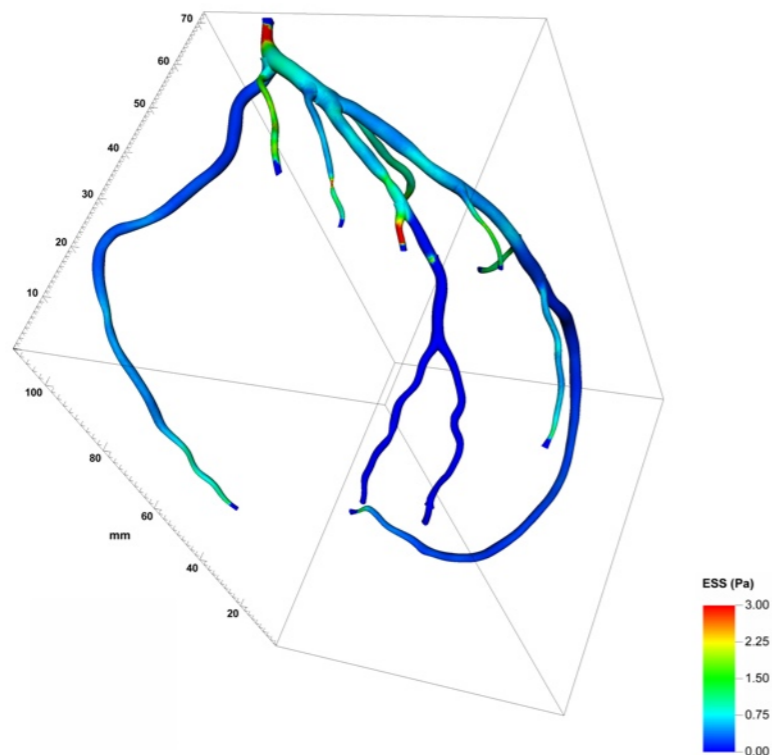
And strong effect of **color**...

39%

percent low ESS regions found

62%

91%



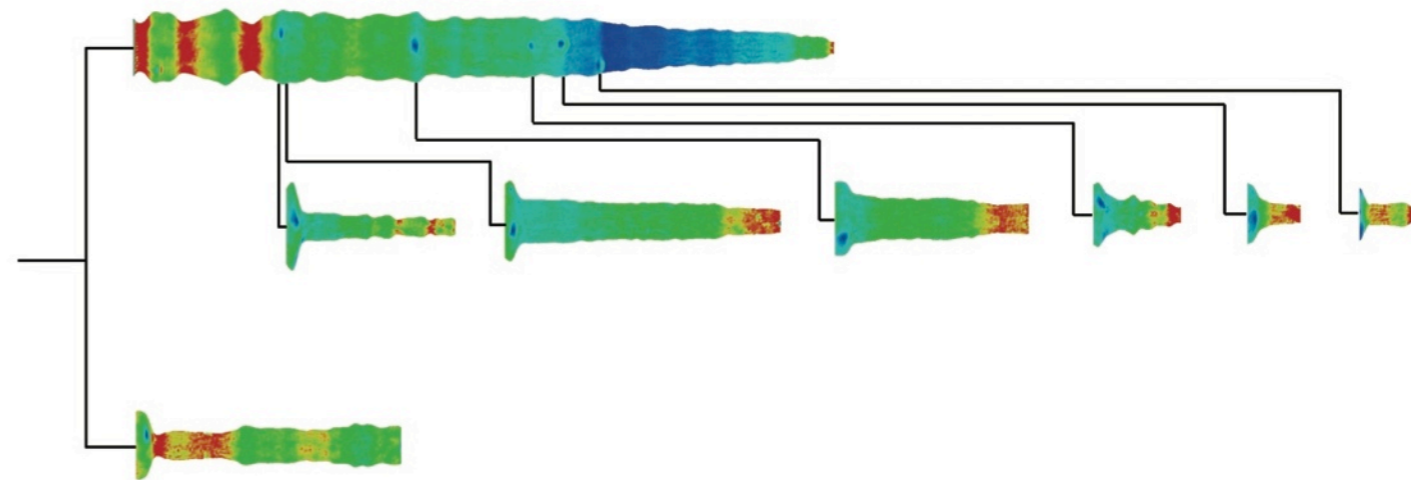
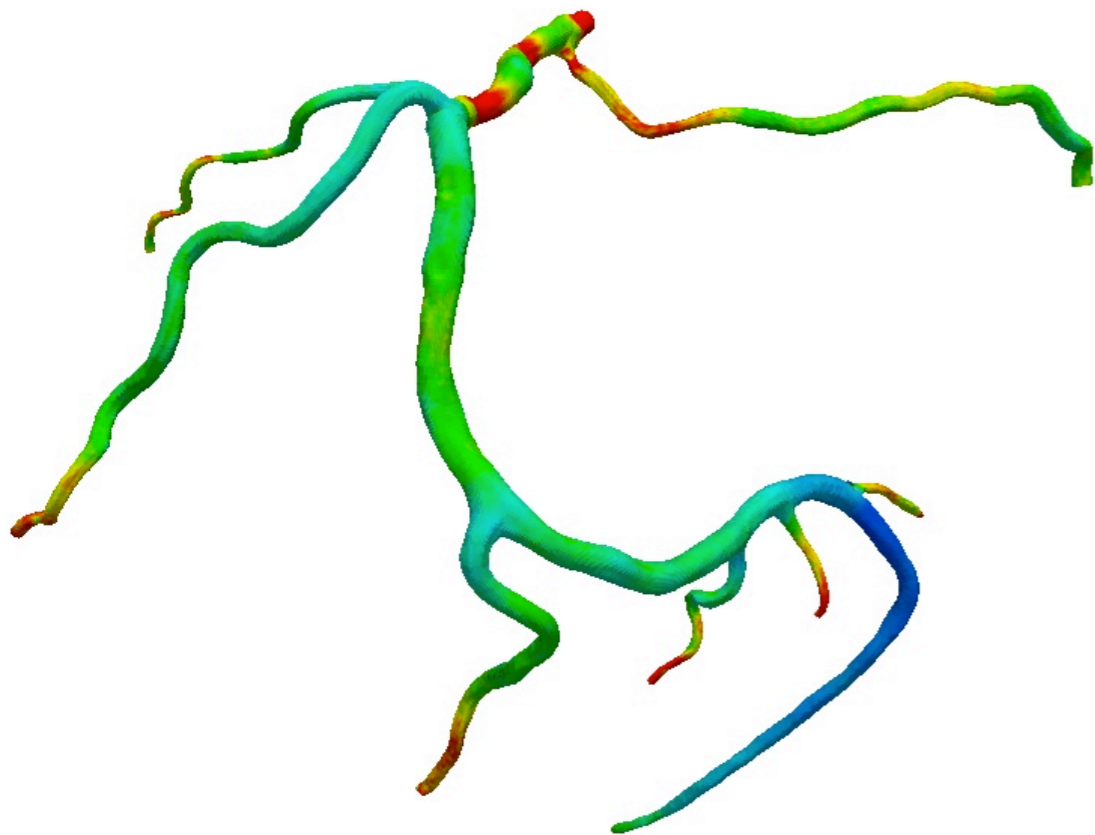
EFFICIENCY

Participants more **efficient** in **2D**.

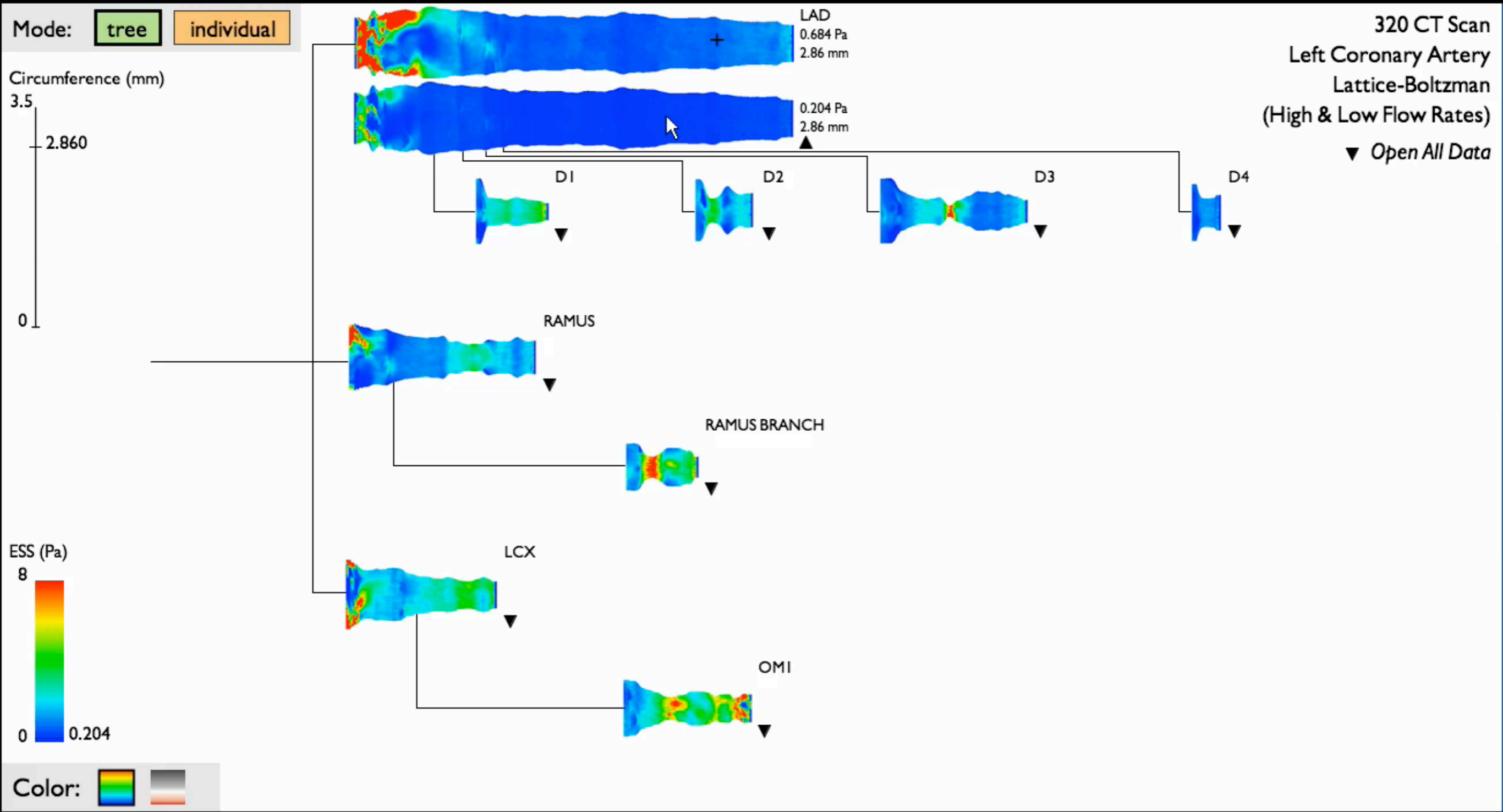
Rainbow color map has greater detriment in 3D.

10.2 sec/region
5.6 sec/region

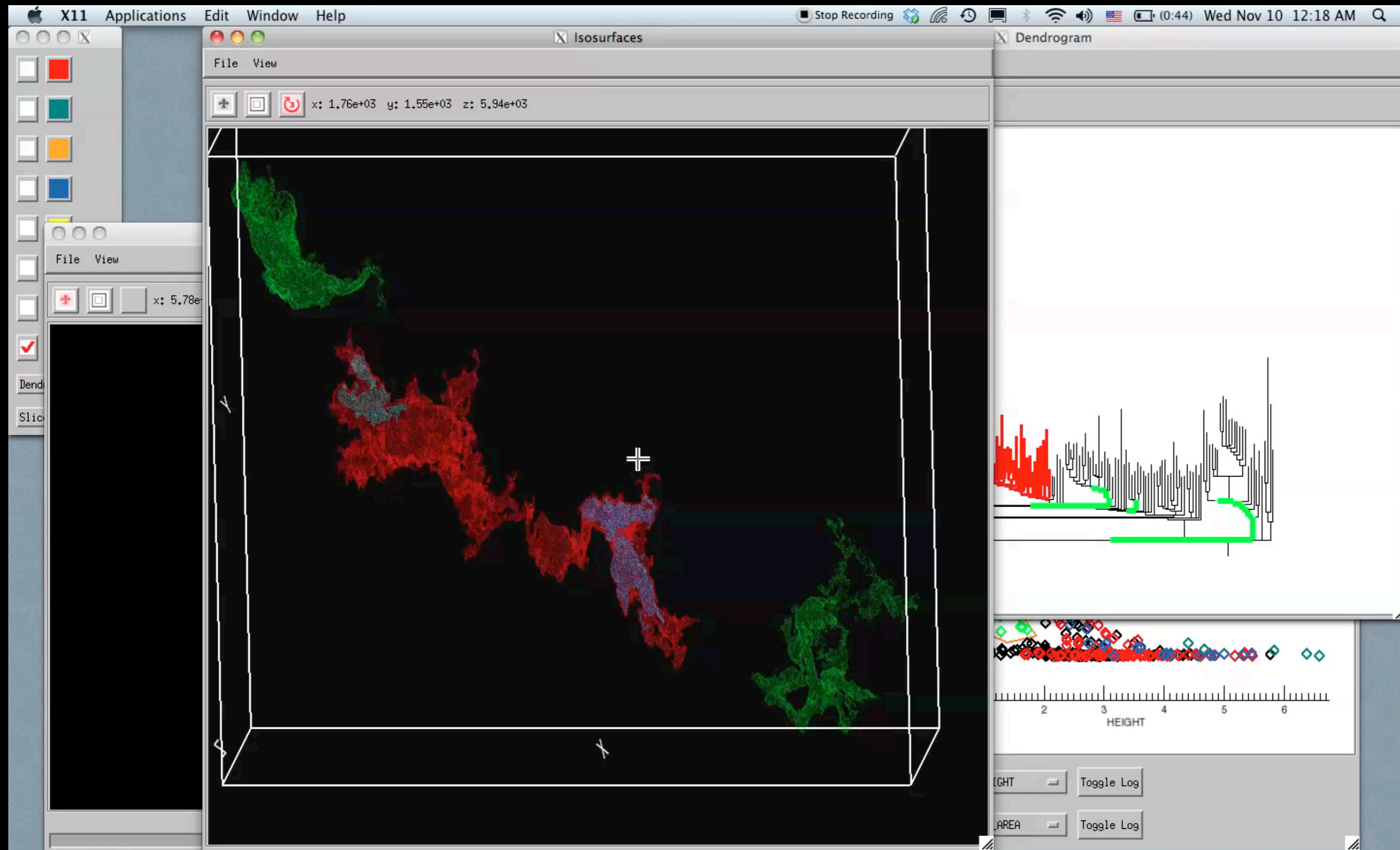
2.6 sec/region
2.4 sec/region



BUT—3D still essential for surgical planning.

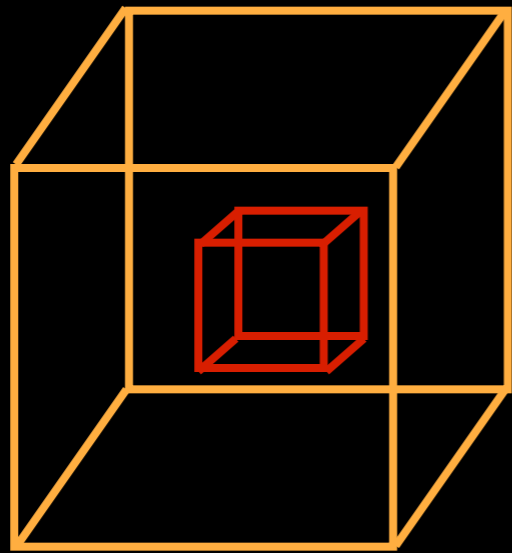


Also in 2011: Linked (Astronomical) Dendrogram Views in IDL

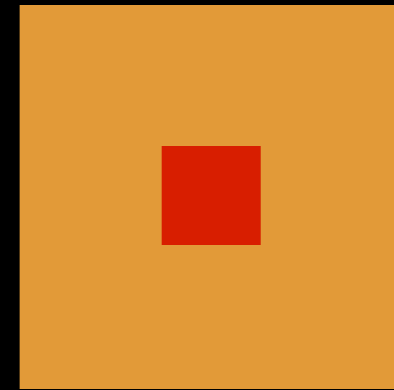


Video & implementation: Christopher **Beaumont**, CfA/UHawaii;
inspired by AstroMed work of Douglas Alan, Michelle Borkin, AG, Michael Halle, Erik Rosolowsky

"Linked Views"

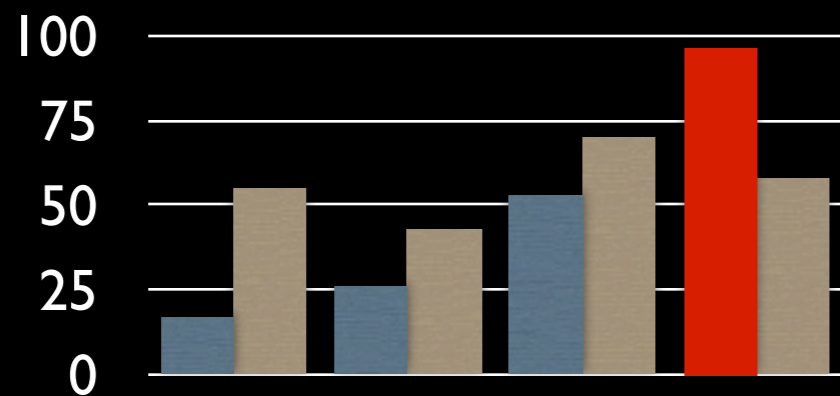


3D

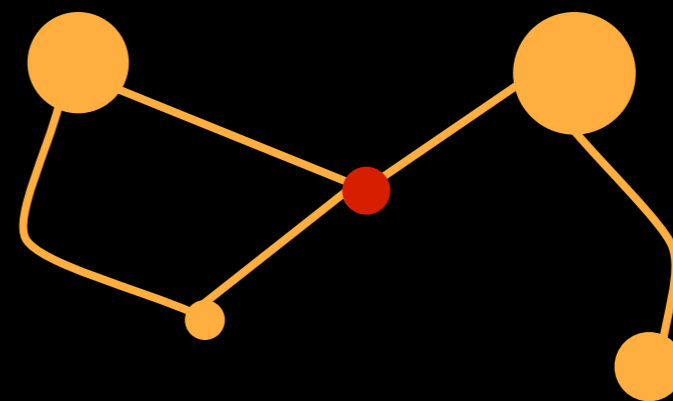


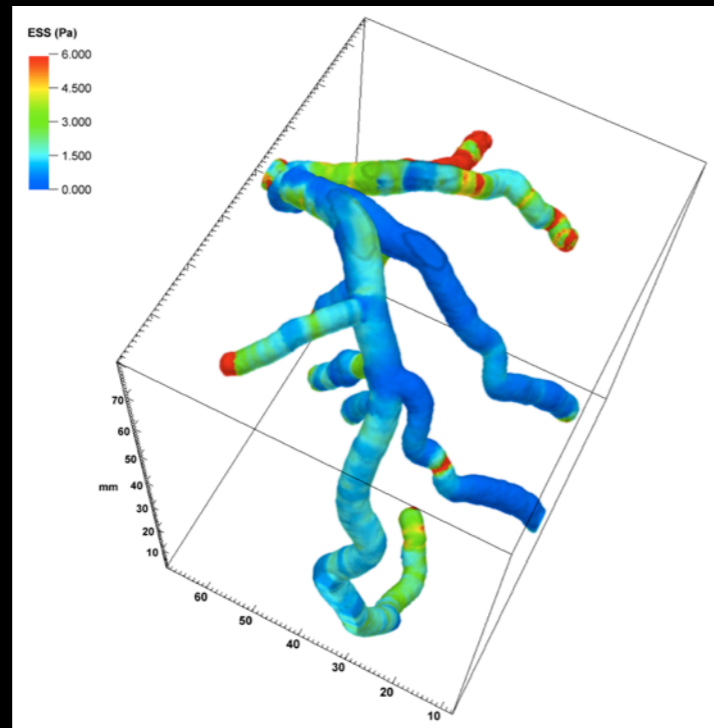
2D

Statistics

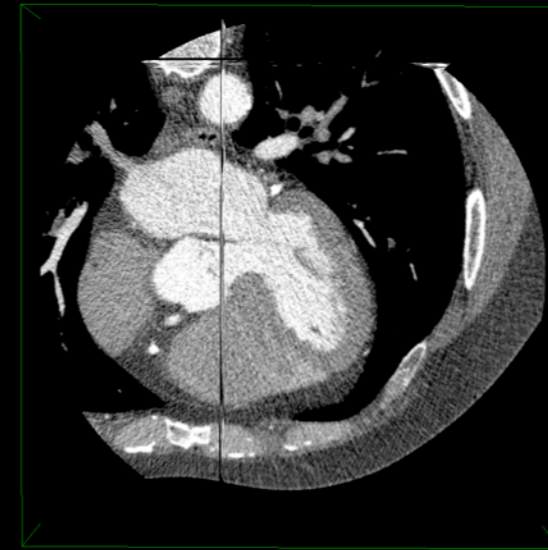


Data Abstraction





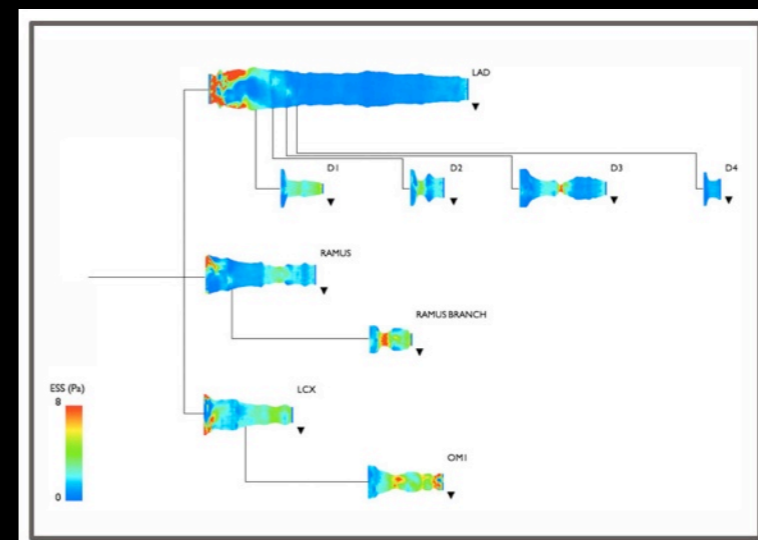
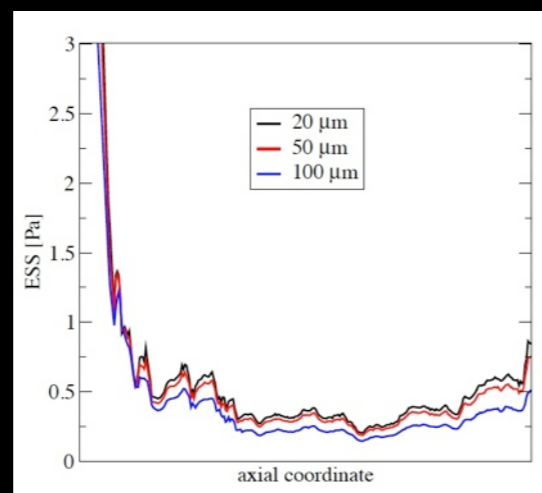
3D

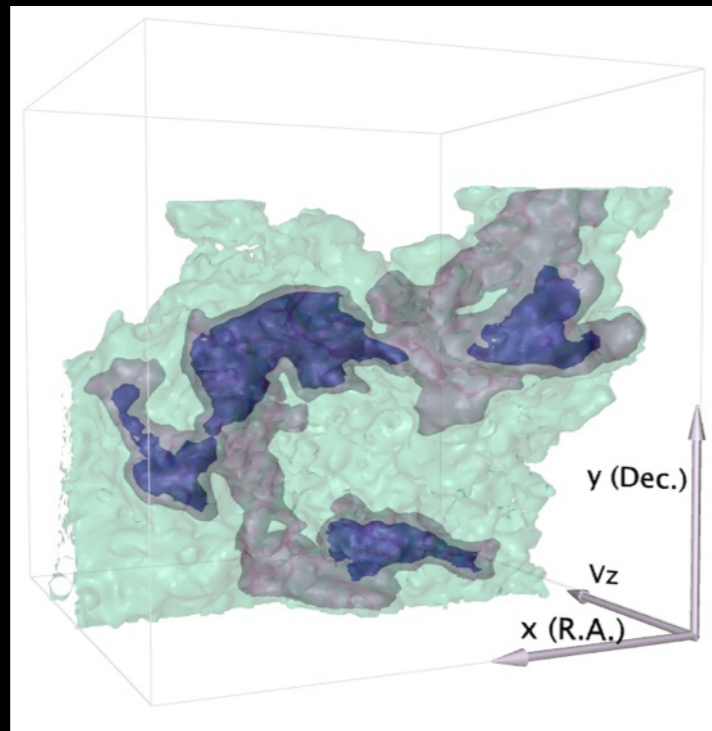


2D

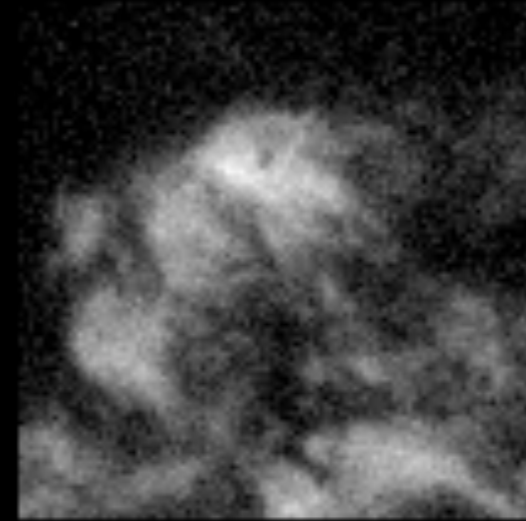
Data Abstraction

Statistics





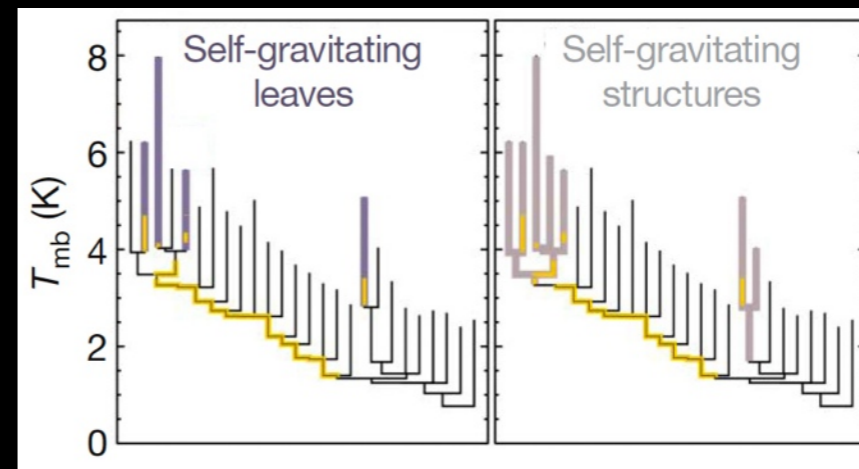
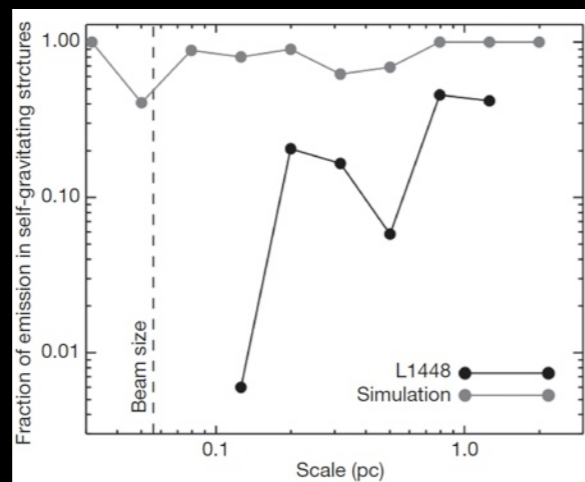
3D



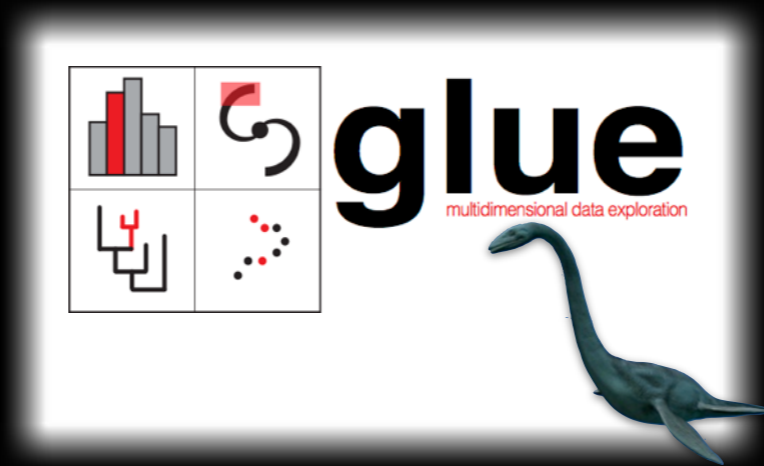
2D

Data Abstraction

Statistics



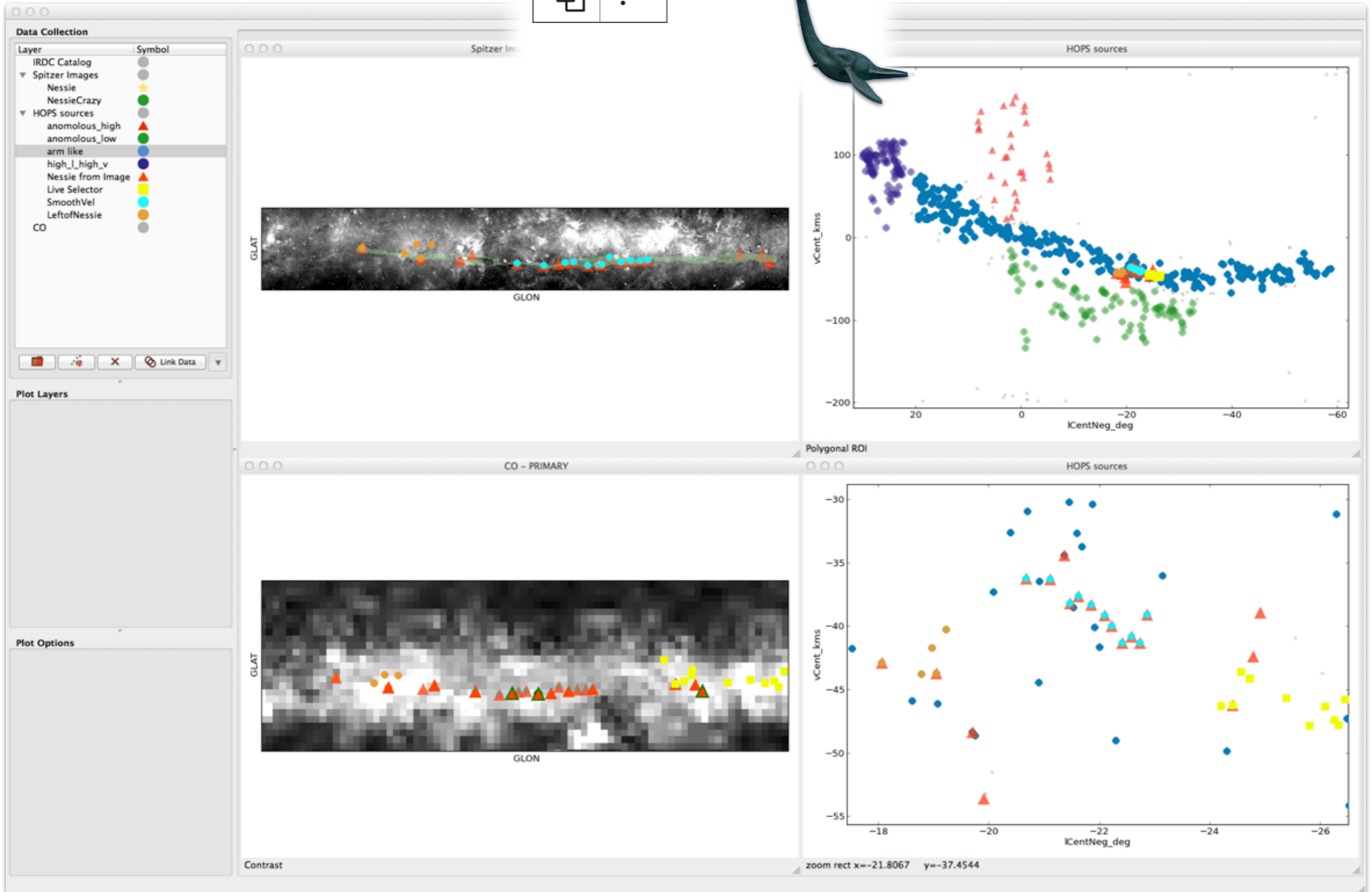
How?





glue


multidimensional data exploration



Glue collaboration: **Beaumont**, Borkin, Goodman, Pfister, Robitaille

What is glue?

Glue 0.1 documentation > next index



Glue Documentation

Table Of Contents
Glue Documentation
Indices and tables

Next topic
Installing Glue

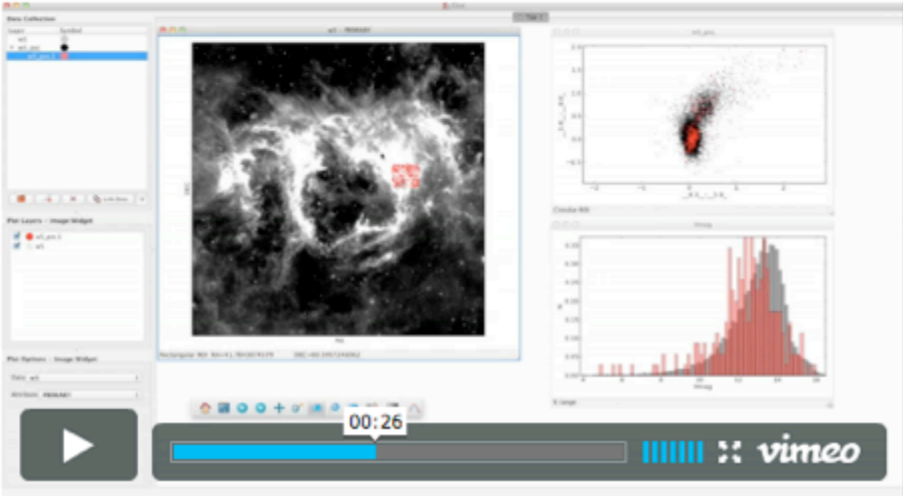
This Page
Show Source
Show on GitHub
Edit on GitHub

Quick search

Enter search terms or a module, class or function name.

Glue is a Python library to explore relationships within and among related datasets. Its main features include:

- **Linked Statistical Graphics.** 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.

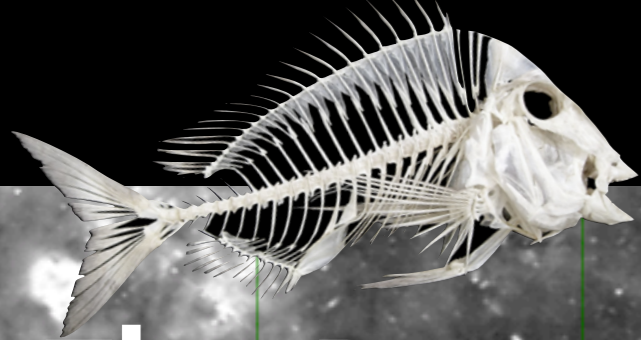


[the film!]

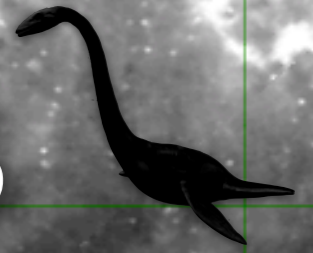
Gluing glue to external APIs

The screenshot displays a video player interface for a demonstration. The main content is a software application window titled 'Python'. On the left, a 'Data Collection' sidebar lists layers: 'IRDC Catalog', 'Spitzer Images', 'HOPS sources', 'HOPS sources.1', and 'CO'. The 'HOPS sources.1' layer is selected. Below this, 'Plot Layers - Image Widget' shows 'HOPS sources.1' and 'Spitzer Images' checked. 'Plot Options - Image Widget' shows 'Data: Spitzer Images' and 'Attribute: IRAC 8.0um'. The central plot, 'Spitzer Images - IRAC 8.0um', shows a colorful astronomical image. A 'Change layer color' dialog box is open over the plot, featuring a color wheel and a 'Density' slider. In the bottom right, a scatter plot titled 'HOPS sources' shows 'Vcent_kms' on the y-axis and 'Position_deg' on the x-axis, with a single red point highlighted. The video player controls at the bottom show a play button, a progress bar at 00:39, and a title 'Glue Demo: World Wide Telescope' from Chris Beaumont, 1 month ago, with a 'NOT YET RATED' badge.

<http://vimeo.com/57078802>



The Bones of the Milky Way: Credits



Seamless Astronomy-style tools used in this project



authorea.com (open publishing)

theastrodata.org (open data)

glueviz.org (open source tools)

universe3d.org (collaborative data)

worldwidetelescope.org (universe information system)

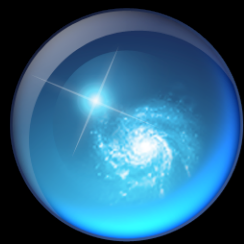
virtual observatory standards (international online information sharing systems)



Supported by



Alyssa Goodman, m:617-230-7080; milkywaybones.org



Microsoft® Research WorldWide Telescope



Experience WWT at worldwidetelescope.org

The screenshot displays the WWT interface with a top navigation bar containing 'Explore', 'Guided Tours', 'Search', 'View', and 'Settings'. Below this is a 'Collections > All-Sky Surveys >' section with thumbnails for 'Digitized Sky Survey', 'VLSS: VLA Low-frequency Sky Survey', 'WMAP ILC 5-Year CMB Temperature Anisotropy Map', 'SFD Dust Map (Infrared)', 'IRIS: Improved Resolution', '2MASS: Two Micron All Sky Survey', and 'Hydrogen Alpha Filter'. The main view shows a large galaxy with a 'Finder Scope' overlay. A 'Context Bar' at the bottom shows 'NGC221' and 'M31'. A 'Context Globe' on the right shows the current field of view on a celestial sphere. A 'Look At' dropdown is set to 'Sky', and a 'Sky' button is highlighted. A 'Finder Scope' window is open, displaying details for 'NGC224'.

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.



John Tukey

Principles of high-dimensional data visualization in astronomy

A.A. Goodman*

Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA

Received 2012 May 3, accepted 2012 May 4

Published online 2012 Jun 15

Key words cosmology: large-scale structure – ISM: clouds – methods: data analysis – techniques: image processing – techniques: radial velocities

Astronomical researchers often think of analysis and visualization as separate tasks. In the case of high-dimensional data sets, though, interactive *exploratory data visualization* can give far more insight than an approach where data processing and statistical analysis are followed, rather than accompanied, by visualization. This paper attempts to chart a course toward “linked view” systems, where multiple views of high-dimensional data sets update live as a researcher selects, highlights, or otherwise manipulates, one of several open views. For example, imagine a researcher looking at a 3D volume visualization of simulated or observed data, and simultaneously viewing statistical displays of the data set’s properties (such as an x - y plot of temperature vs. velocity, or a histogram of vorticities). Then, imagine that when the researcher selects an interesting group of points in any one of these displays, that the same points become a highlighted subset in all other open displays. Selections can be graphical or algorithmic, and they can be combined, and saved. For tabular (ASCII) data, this kind of analysis has long been possible, even though it has been under-used in astronomy. The bigger issue for astronomy and other “high-dimensional” fields, though, is that no extant system allows for full integration of images and data cubes within a linked-view environment. The paper concludes its history and analysis of the present situation with suggestions that look toward cooperatively-developed open-source modular software as a way to create an evolving, flexible, high-dimensional, linked-view visualization environment useful in astrophysical research.