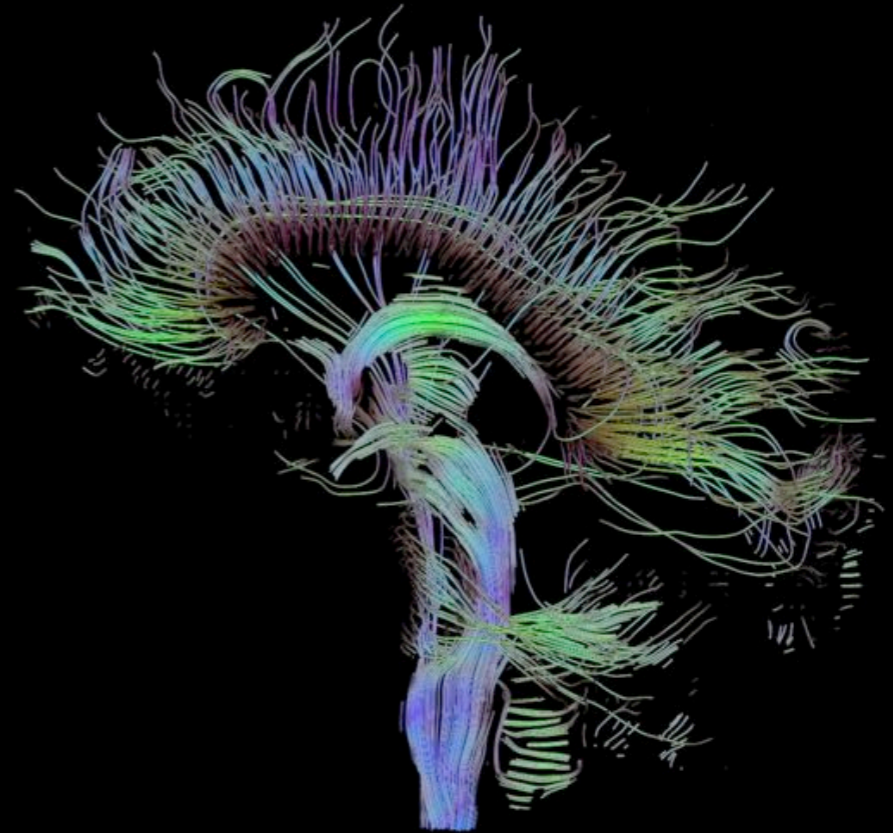
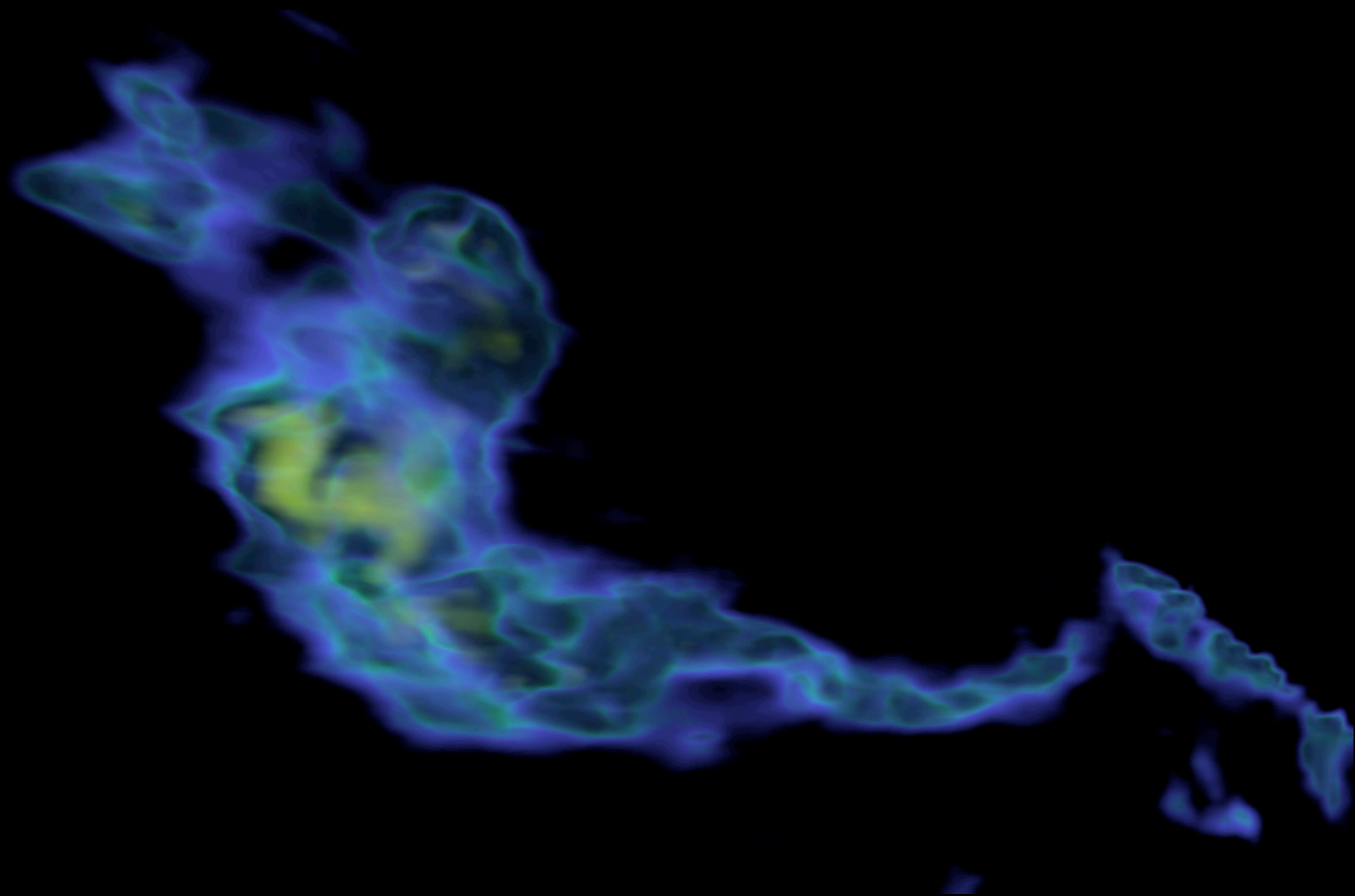


# Data Intensive Science in Astronomy & Medicine



Alyssa A. Goodman  
Initiative in Innovative Computing @ Harvard  
and



Harvard-Smithsonian Center for Astrophysics

# Shared Challenges...Shared Approaches?



Astronomy

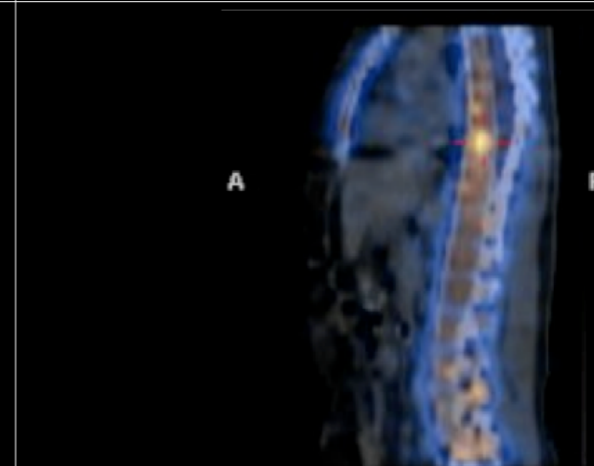
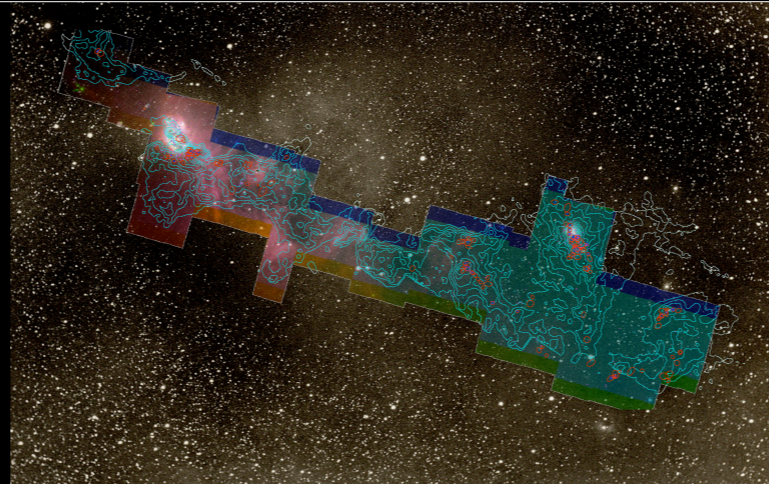


Medicine

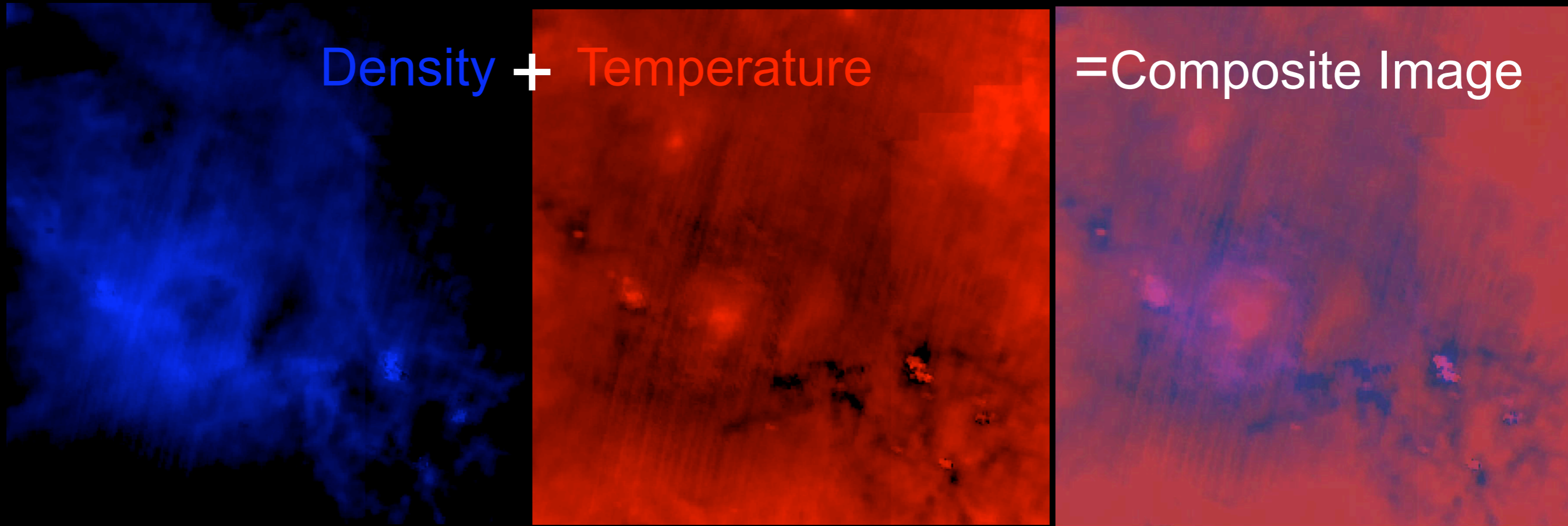
# Shared Challenges...Shared Approaches?



Heavily Reliant on  
"Multi-Modal Imaging"

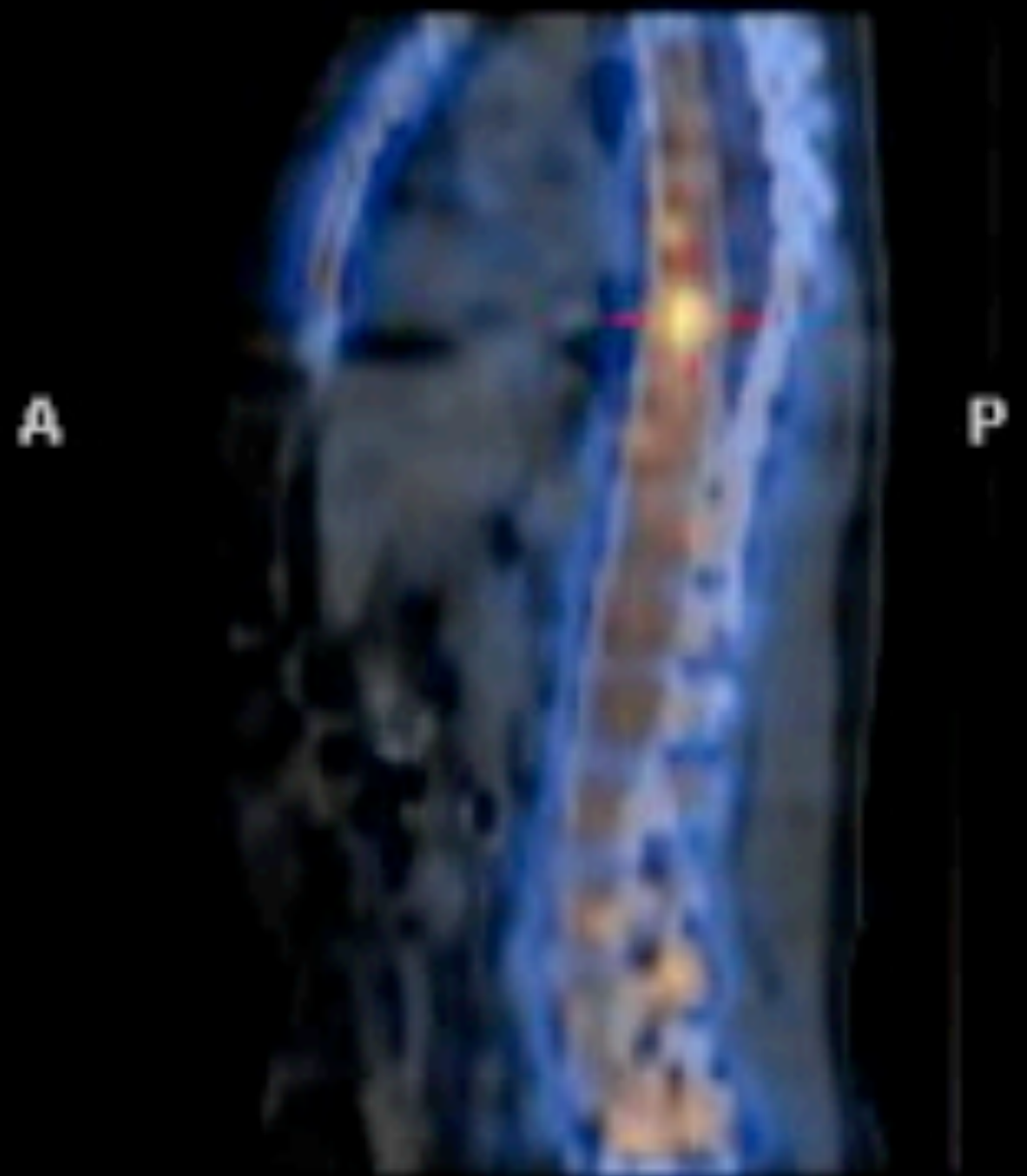
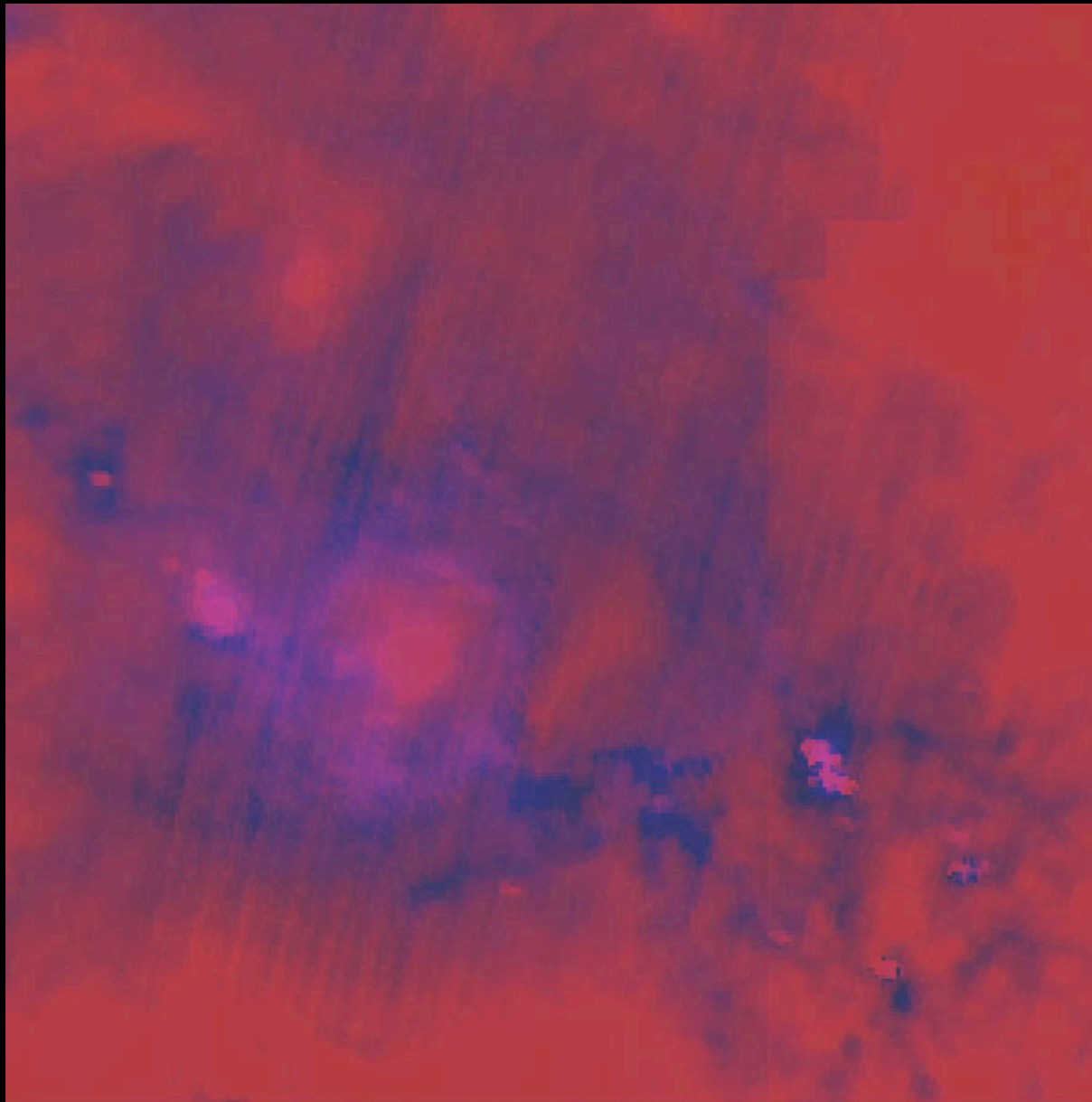


# “Multimodal Imaging”



*Dust mapping of Perseus based on IRAS data included in the COMPLETE Survey of Star-Forming Regions  
(Schnee et al. 2005; Ridge et al. 2006)*

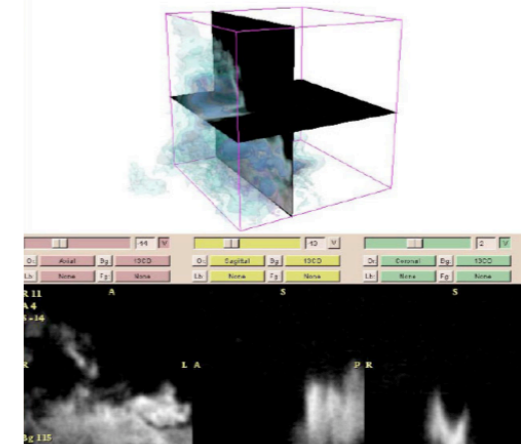
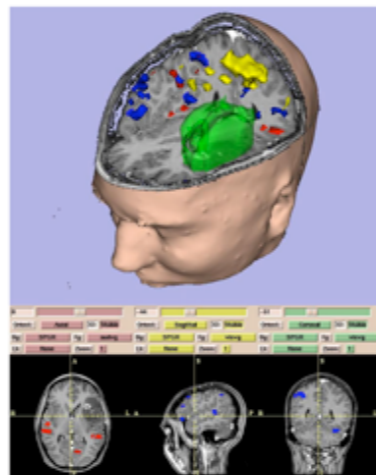
# Astro-Medical Multimodal Imaging








# Shared Challenges...Shared Approaches?



Increasingly reliant on 3D  
and 4D imaging



# COMPLETE = COordinated Molecular Probe Line Exinction Thermal Emission

-  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. in prep.)
-  Optical image (Barnard 1927)

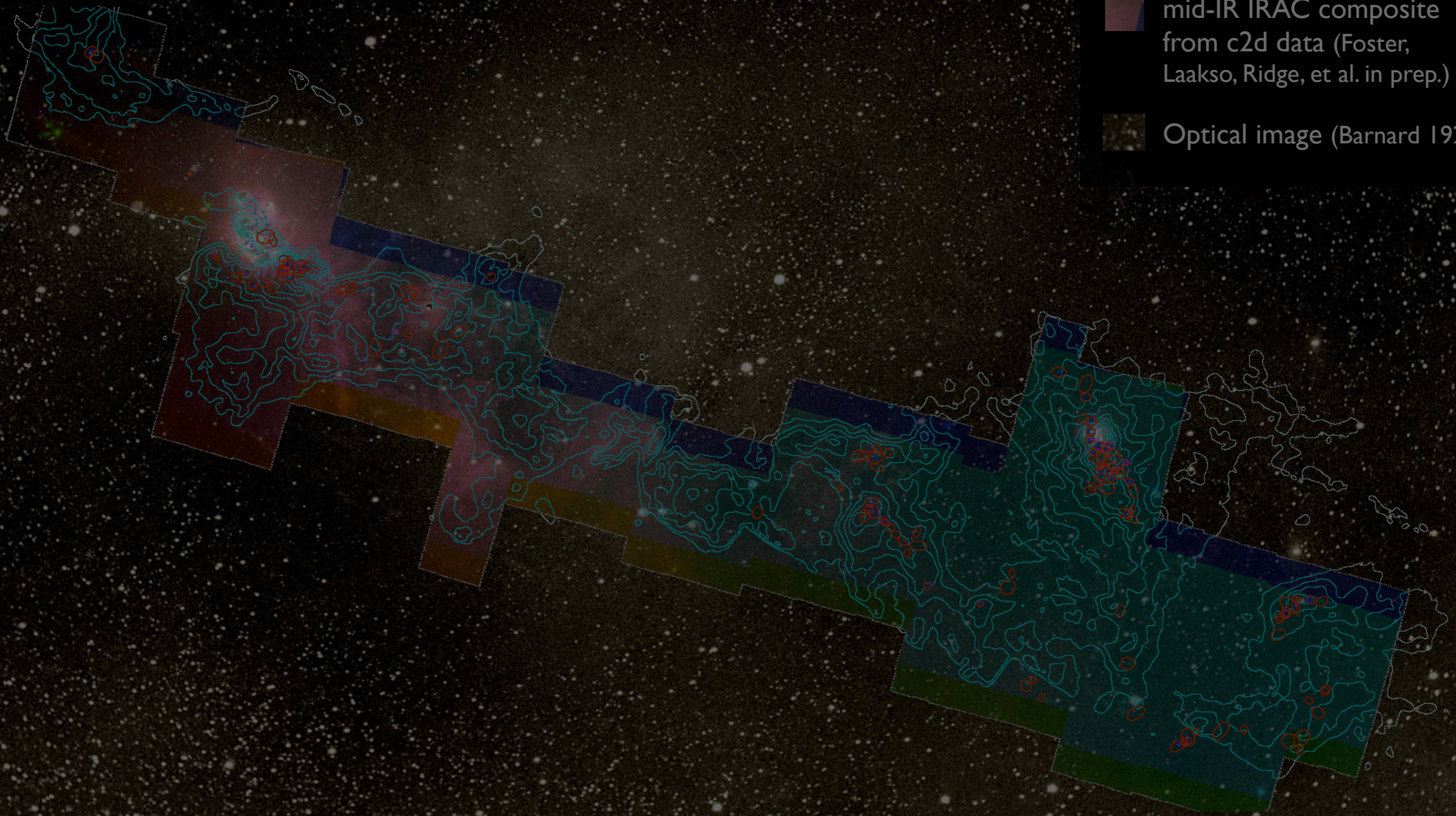
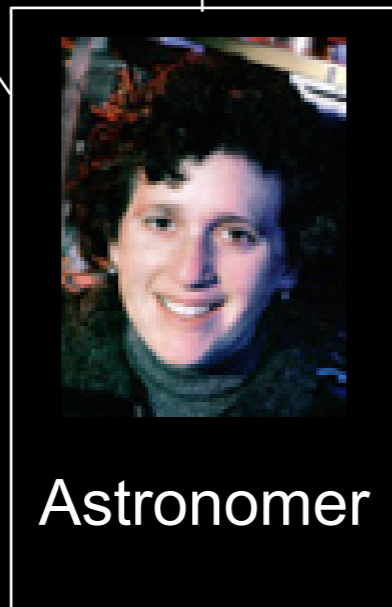
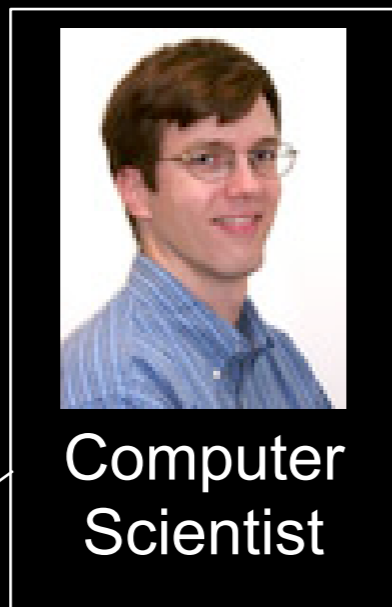
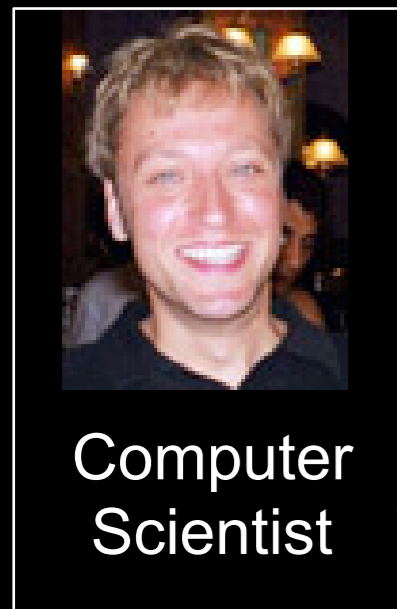


image size: 520 x 274  
view size: 1305 x 733  
URL: 63 WWW

m: 17249  
zoom: 227% Angle: 0

# The Astronomical Medicine Story



*“Viz has failed the scientific community...”*

The screenshot shows the IIC website header with the logo and the text "Initiative in Innovative Computing at Harvard". Below the header is a section titled "projects" with a sub-section for "Astronomical Medicine". This section features two images: a 3D visualization of a nebula within a wireframe cube and a 2D heatmap of a similar nebula. Below the images are sections for "Lead Investigators" (listing Alyssa Goodman, Mike Halle, Ron Kikinis, and David Kennedy), "Project Staff" (listing Doug Alan, Michelle Borkin, and Jens Kauffmann), and a "Description" paragraph.

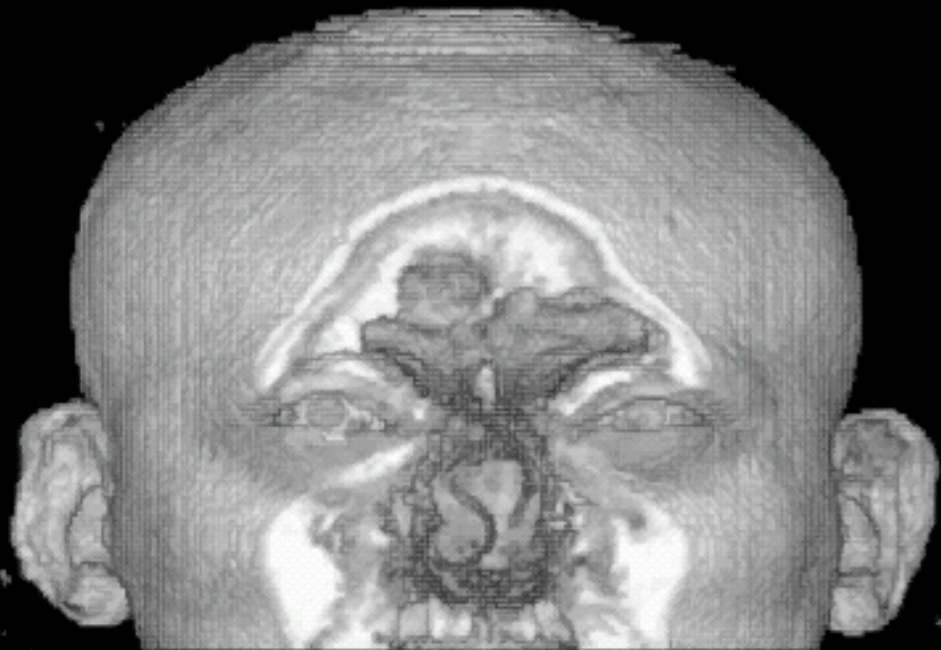
- +Nick Holliman (CS, 3D expert)
- +Doug Alan (S/W Engineer)
- +Jens Kauffmann (postdoc)





# "Slices"

"KEITH"



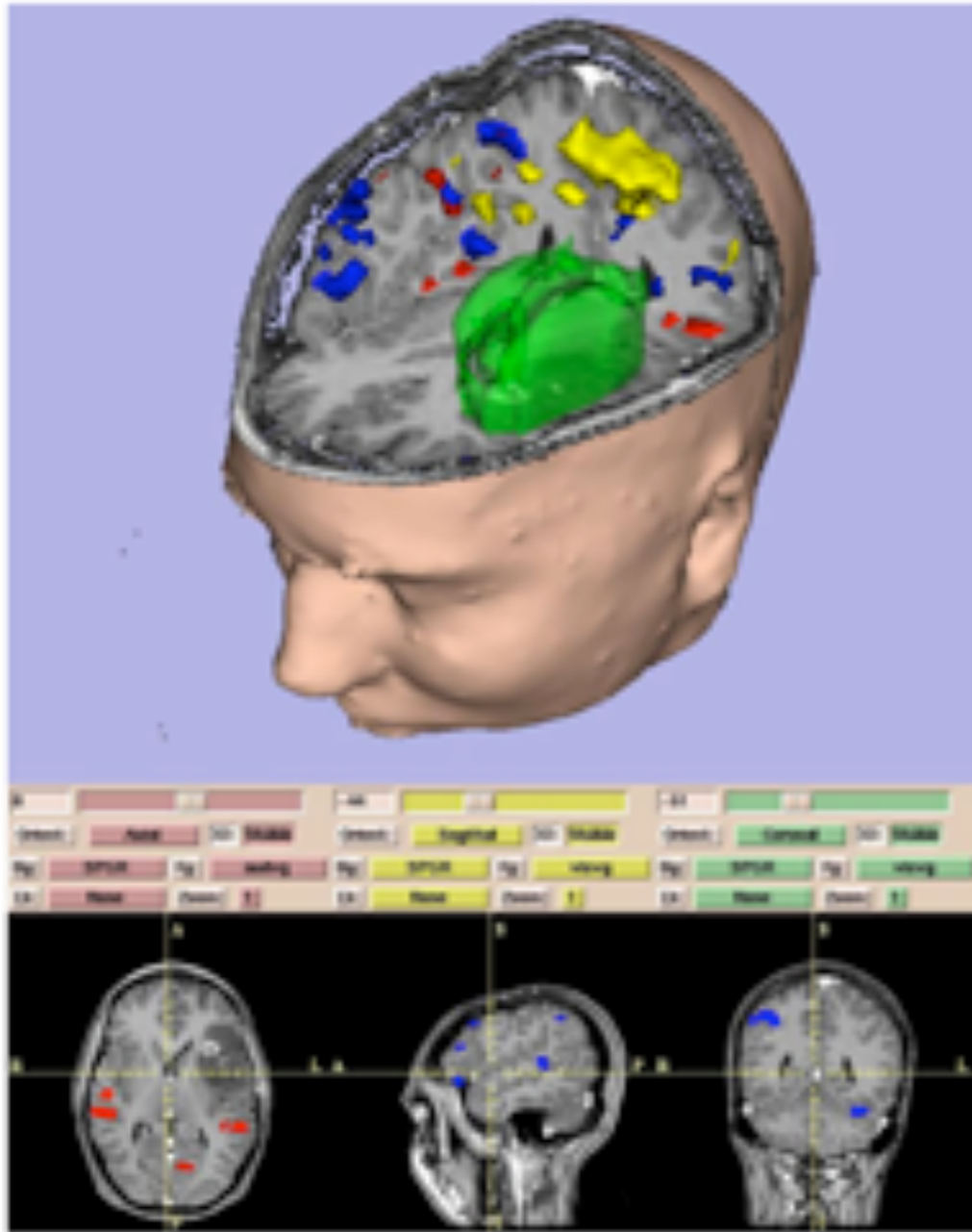
"PERSEUS"



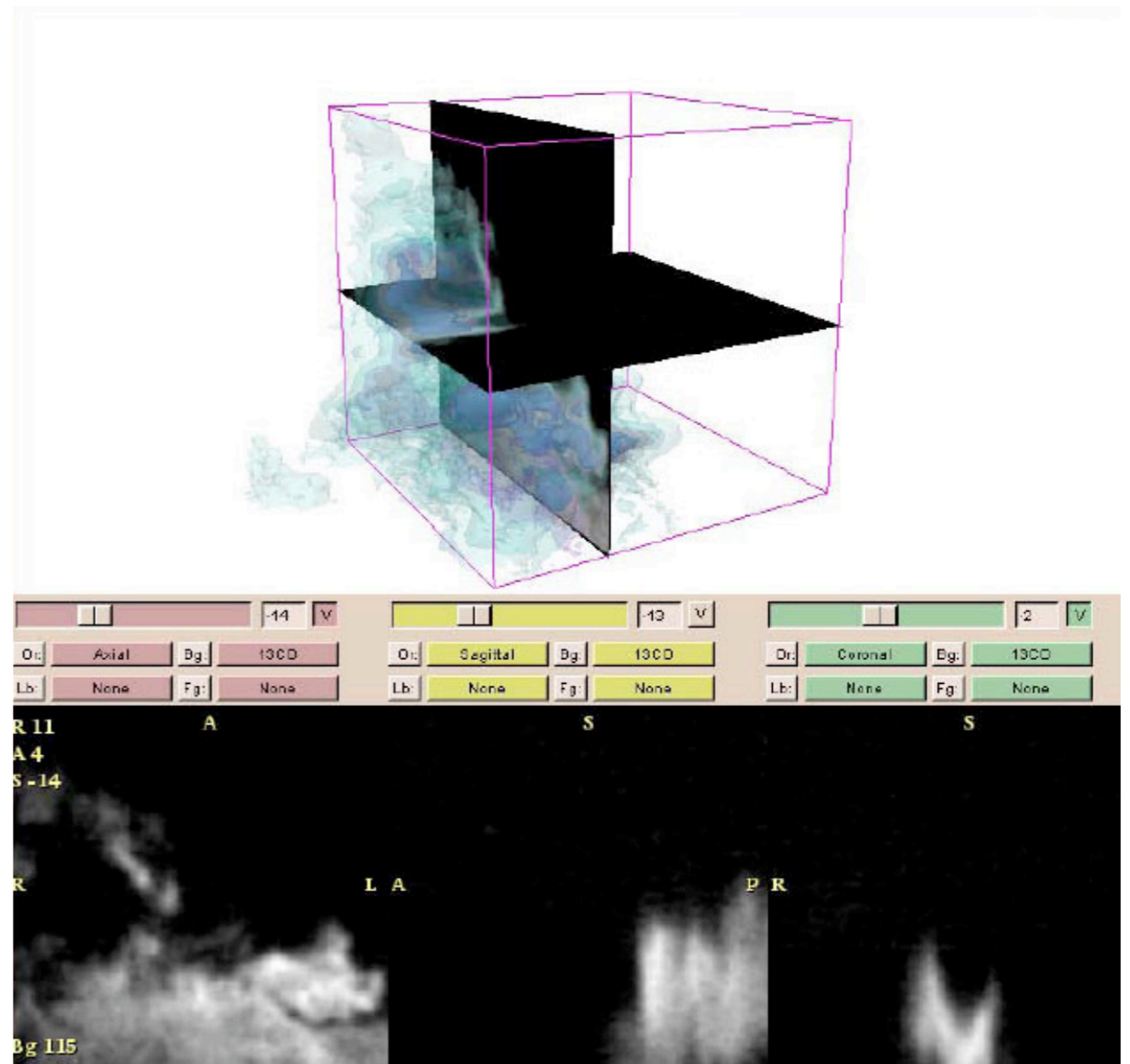
"z" is depth into head

"z" is line-of-sight velocity

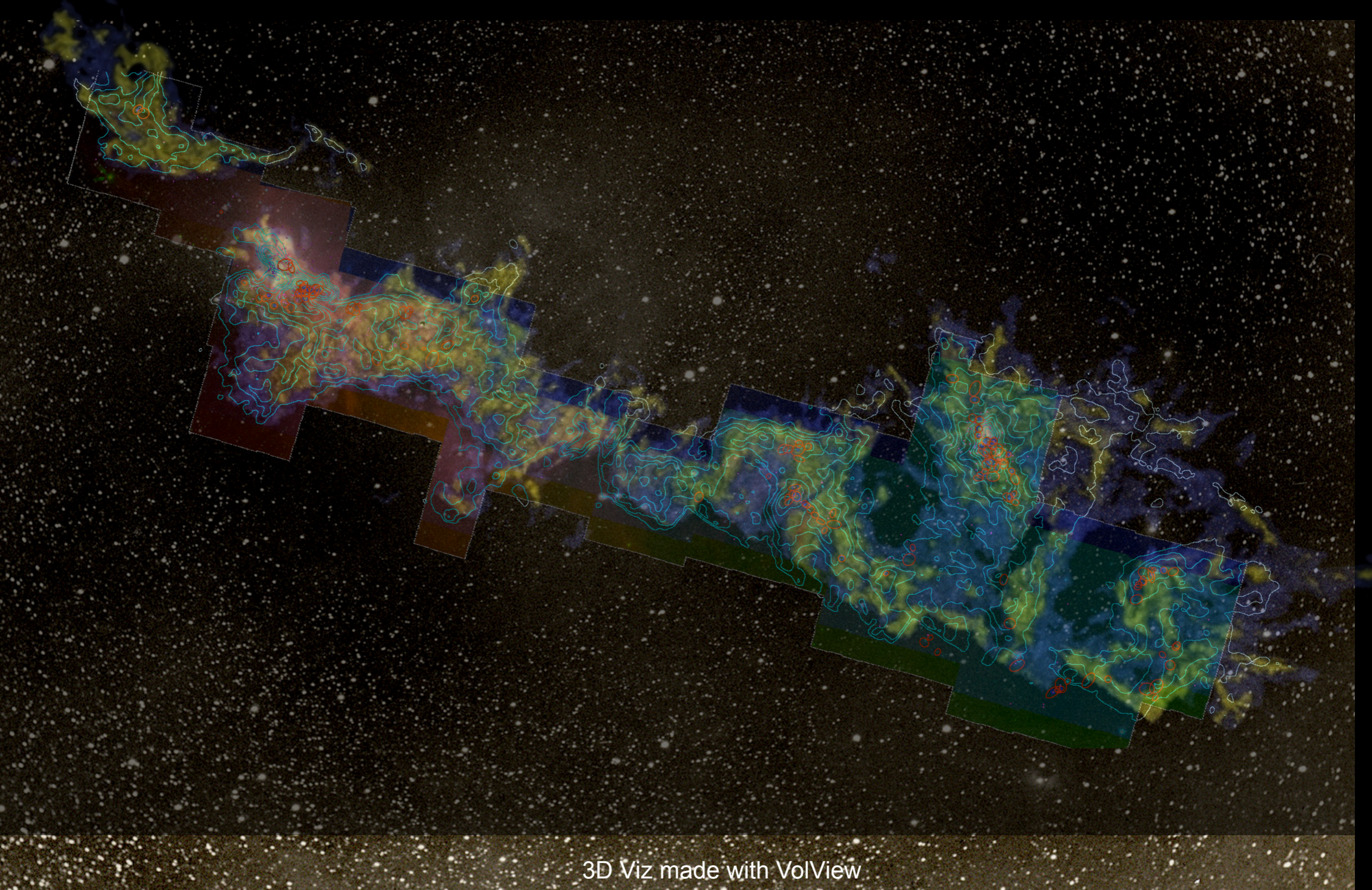
# “Astronomical Medicine”



MRI



Spectral-Line Data Cubes



3D Viz made with VolView

# Shared Challenges...Shared Approaches?



Astronomy



Medicine

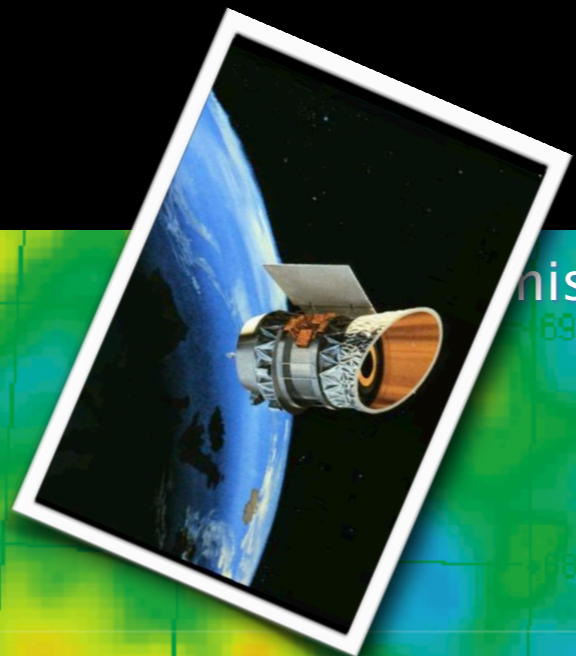
**Distributed Users, with  
needs to combine diverse,  
distributed, data sets**

“The Virtual Observatory”

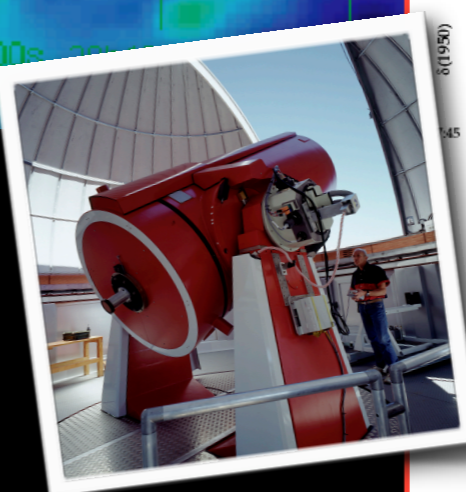
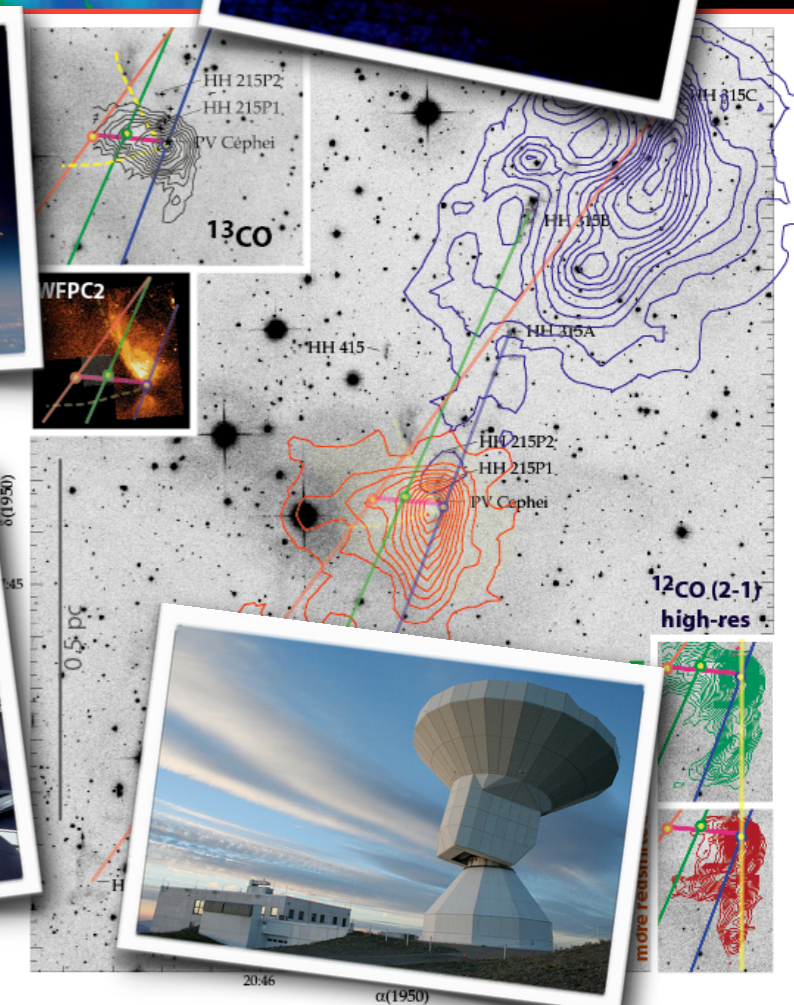
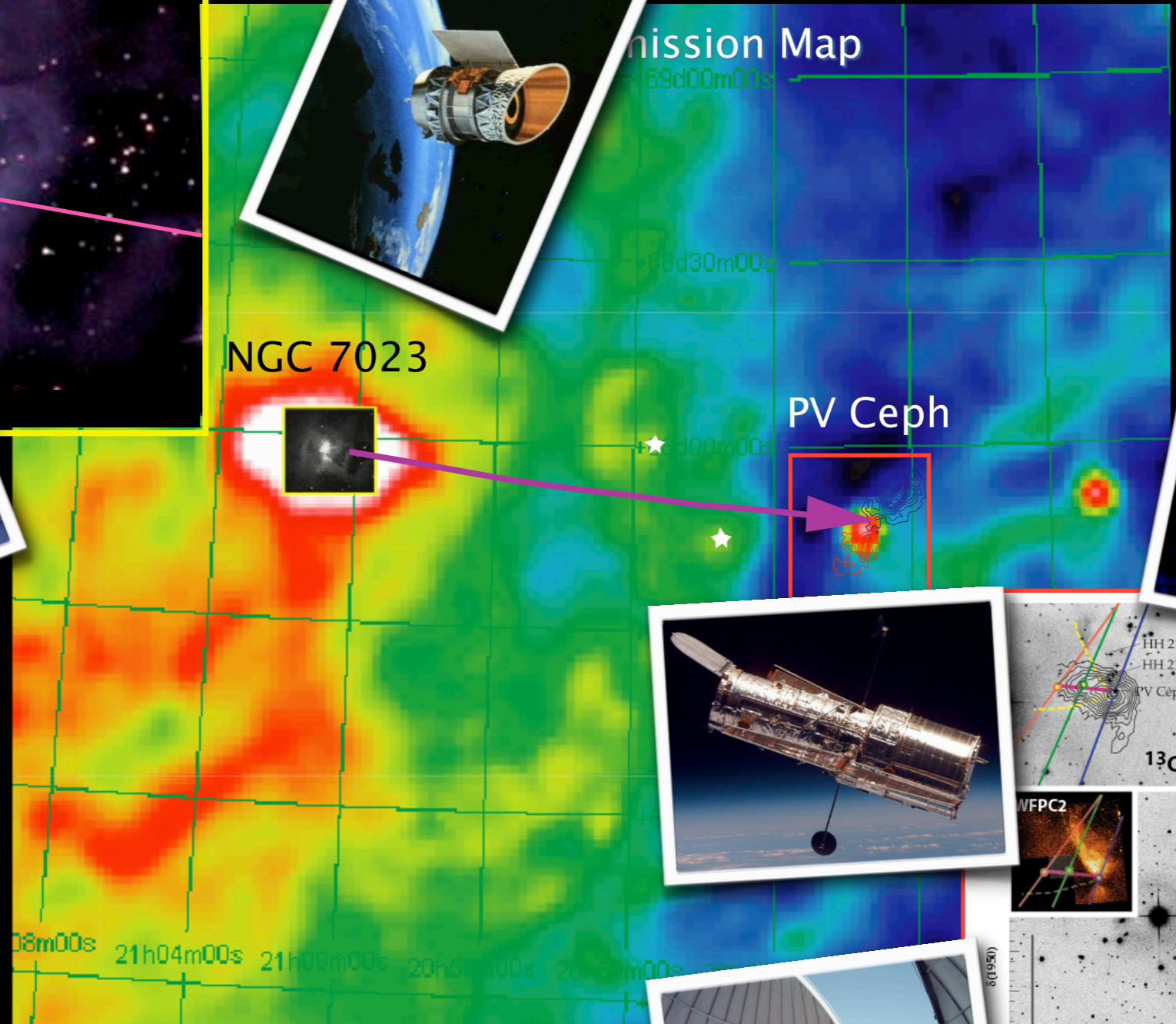
“The BIRN” (Biomedical  
Informatics Research  
Network)

# Optical Image of NGC 7023

0.5 pc

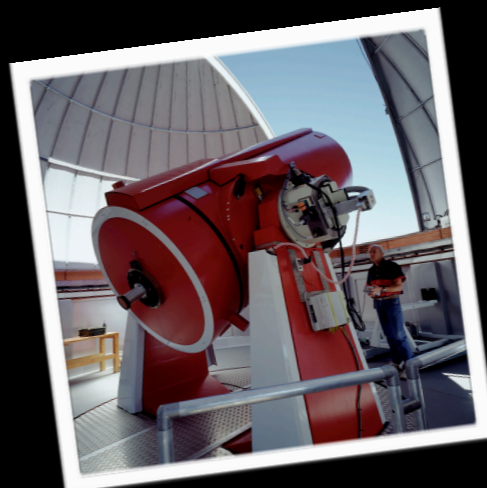
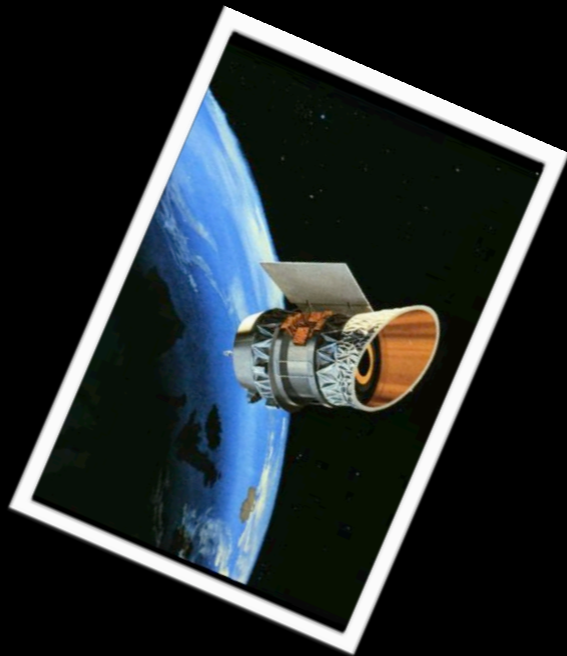


mission Map



# Star Running Away from Home!

Goodman & Arce 2004



Six *Different* Telescopes...  
and how did we get the data??

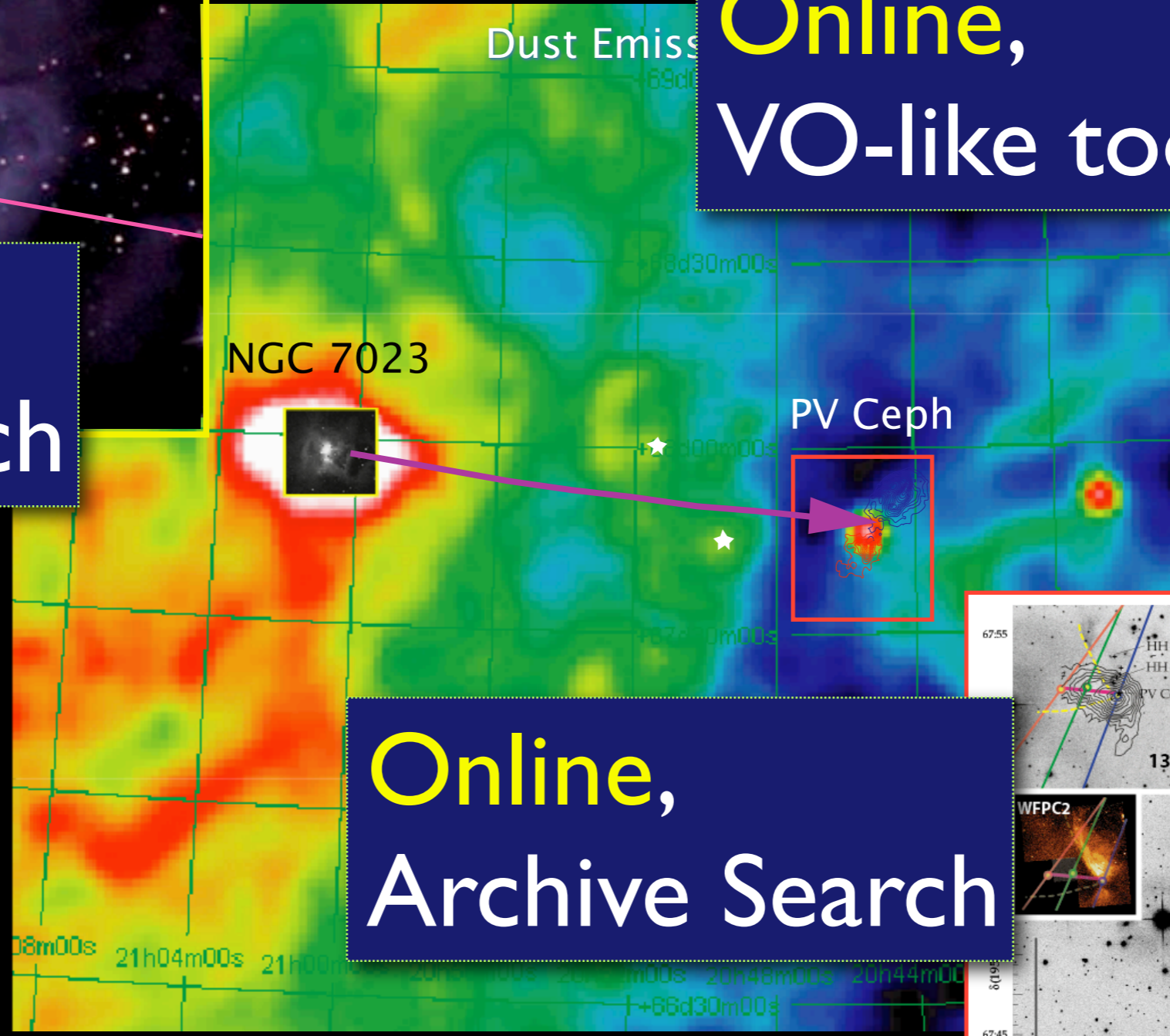


# Optical Image of NGC 7023

0.5 pc

Online,  
Web Search

Online,  
VO-like tool

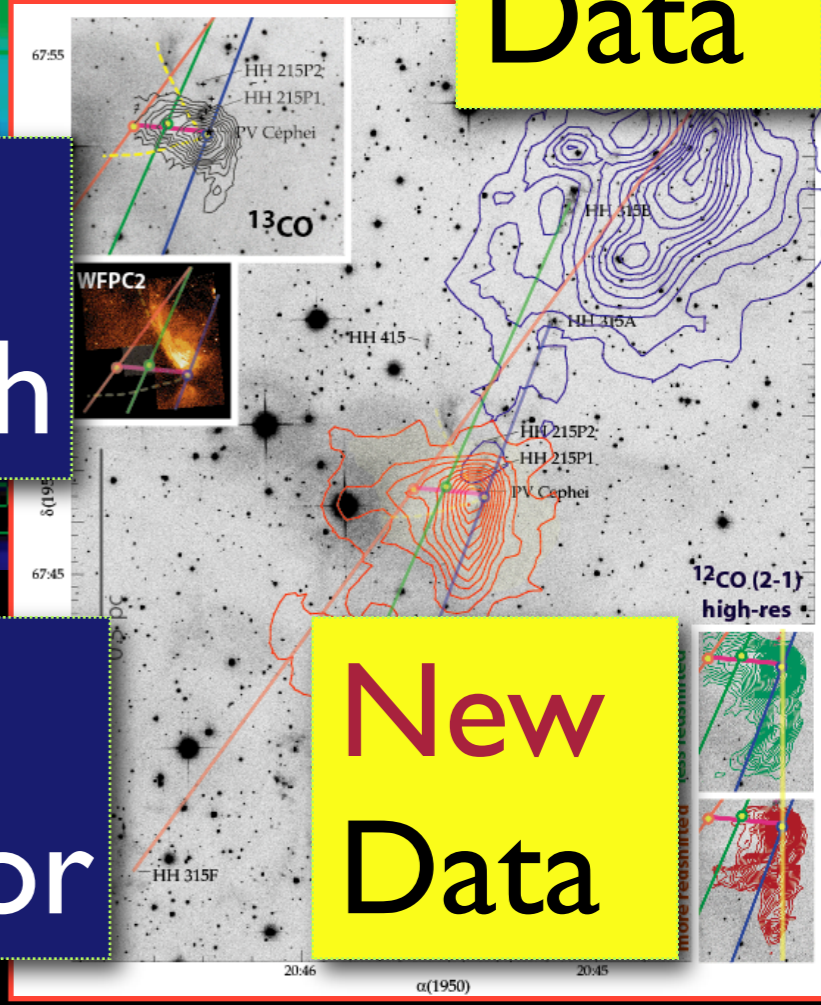


Online,  
Archive Search

New  
Data

Email  
from Author

New  
Data



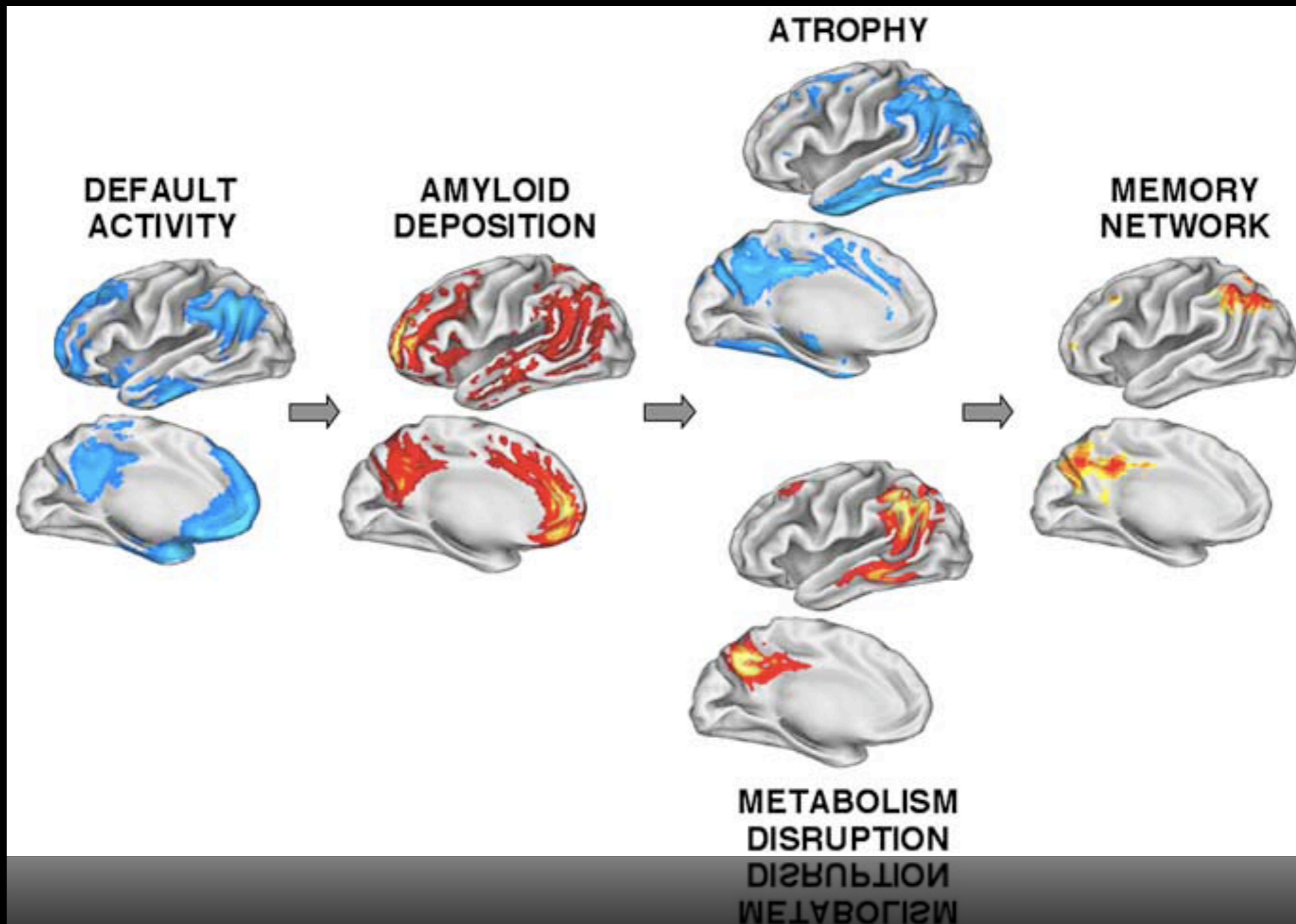
# What do I want?

Seamless tools that let me make discoveries like this, for:  
any object, or set of objects, in the sky  
...in seconds

...and then I would like to share what I've done with  
others, through publications and *much* more.



...and my friends want that too...



e.g. **“Neuroinformatics Framework”** at IIC (Buckner, Clark, Fariello, Hyman, Rosen, et al.); image from Buckner et al. 2005,

# Publishing & Sharing

The screenshot shows a PDF viewer window titled 'astromed\_3d\_paper.pdf'. The interface includes a top toolbar with options like 'Create PDF', 'Combine Files', 'Export', 'Secure', 'Sign', 'Forms', and 'Review & Comment'. Below the toolbar is a navigation bar with page number '1 / 1', a zoom level of '66.7%', and a 'Find' search box. On the left side, there is a 'Model Tree' sidebar with a 'Highlight Color' dropdown and 'Options' menu. The tree structure shows a hierarchy: 'node #-1' containing 'Group', 'outer contour', 'Drawings', and 'Vectorized Layer'. Below the tree is a 'Property' panel with a list of attributes: 'Number of Lights', 'Number of Children', 'Number of Vertices', 'Number of Faces', 'Number of Textures', and 'Bounding Box Dimensions (Mete)'. The main content area displays a scientific paper with text, a graph, and a map. The text discusses star formation and dendrograms. The graph shows 'Relative Density' vs 'x' with 'Dendrogram Branches' highlighted. The map shows 'Isointensity surface models of L1448 in <sup>13</sup>CO'. A caption below the map reads: 'Figure 2 | Isointensity surface models of L1448 in <sup>13</sup>CO. The models are in p-p-v space, where the front of the cube is the plane of the sky. The intensity thresholds shown are chosen using the "dendrogram" procedure described.'

structures, but also many odd features corresponding to dimples on the surfaces of larger ones.

So, in a simple picture of star formation, where one "clump" lasts a long time, and forms one star, it makes sense to think about breaking up the cloud into non-overlapping volumes and creating a "clump mass spectrum" from the result. But, in a turbulent picture, where hierarchical structure pervades the cloud and structures are transient, it seems foolhardy to break up the full volume of the cloud into non-overlapping clumps. The analysis is further complicated by the aforementioned 2 spatial+1 velocity dimension nature of the observational data. Segmentation based analysis algorithms typically ignore this subtlety and operate in three dimensions despite the fact that much of the clumpy structure identified is thought to result from chance superpositions along the line of sight (Ostriker et al.). We note that the segmentation of sparse two-dimensional maps via the CLIMPFIND is not as fraught since the standard application of the algorithm in this case is to relatively unblended data for source identification (similar to the standard Source EXtractor algorithm Bertin, E.; Amouts, S. A&AS 1996).

In this letter, we borrow heavily on techniques used in other fields to show that a novel application of commonly used structure trees (e.g. refs. from NSF proposal) to molecular line data provides a method to characterize the hierarchical structure we use the results to identify physically relevant features in the data. While well-developed in other fields such as computer science and computational biology, the application of the tree methodology in astrophysics has been relatively lacking. In cosmological simulations, the merger history of galaxies is frequently parameterized as a function of redshift with a structure tree (e.g. Kauffmann & White, MNRAS 1993). In the field of star formation, Houlihan & Scalzo (1990) proposed applying structure trees to extinction maps to characterize their hierarchical structure. Using this as an inspiration, we developed an algorithm that "abstracts" the hierarchical structure of a p-p-v data cube into an easily visualized representation; and then we use this abstraction to identify the structures relevant to star formation in the data.

**DENDROGRAMS**

A schematic illustration of the dendrogram process appears, showing the construction of a dendrogram from a hypothetical 1D emission profile (blue). The "dendrogram" (in black) graphically represents the peaks of the emission structure as separate branches. At the highest  $T_{mb}$  value that still spans both peaks, the two branches are connected into a single branch. By repeating this process for every peak of the emission, dendrograms abstract a complete topological description of the emission in a data cube<sup>1</sup>. We construct analogous dendrogram diagrams for 3D data cubes; however, to plot them most clearly in 2D, we flatten them, eliminating any meaningful spatial information on the x-axis.

Our new contribution to the dendrogram technique is the calculation physical properties for every object corresponding to a branch in the dendrogram<sup>2</sup>. To determine what features might be important, we calculate a virial parameter as the ratio of kinetic

energy to gravitational binding energy (without external pressures or a magnetic field) for every branch in the dendrogram. We highlight in red every branch that corresponds, in our simplified model, to a self-gravitating object. The standard feature identification algorithms may find the objects at the top branches of the dendrogram trees, since these correspond to the peaks of "clumps," but would be unable to identify objects at the base of the dendrogram "tree."

In Figure 1 we show the dendrogram of the L1448 region in the star-forming molecular cloud Perseus. The original data (Figure 2) are <sup>13</sup>CO(1-0) emission taken from the COMPLETE survey of Star Forming Regions survey of the Perseus region (Ridge et al.). The main complex in L1448, represented by the dominant branch of the dendrogram, has gravitational and kinetic energies that are comparable and encompasses a large fraction of the region.

The dendrogram indicates the importance of gravity over a large range of scales in the molecular cloud from the individual small clumps indicated with the "leaves" of the dendrogram down into the base of the structure tree. In contrast, the dendrogram indicates that the feature found at large velocities is dynamically as well as kinematically distinct from the majority of the region. In addition, the dendrogram identifies several "leaves" which have distinct regions where gravity dominates on the smallest scales. Such features are interesting since they represent where the molecular gas is closest to forming stars.

<sup>1</sup> Dendrogram construction is completely determined by the data without relying on choice algorithm parameters, such as the step size in the CLIMPFIND algorithm or the relative weights of the c2 components in Gasclouds.

4

Goodman et al., *Nature*, submitted June 2007;  
see also Barnes & Fluke 2007 arXiv:0709.2734

# Shared Challenges...Shared Approaches?



Astronomy



Medicine

**Distributed Users, with  
needs to combine diverse,  
distributed, data sets**

“The Virtual Observatory”

“The BIRN” (Biomedical  
Informatics Research  
Network)

# What is “the” Virtual Observatory?

*(see Alex Szalay's talk  
in this Symposium)*

Google™

virtual observatory

Search

[Advanced Search](#)  
[Preferences](#)

Web

## [SkyView Virtual Observatory](#)

**Virtual observatory** on the net generating images of any part of the sky at wavelengths in all regimes from radio to gamma-ray.

[skyview.gsfc.nasa.gov/](http://skyview.gsfc.nasa.gov/) - 10k - [Cached](#) - [Similar pages](#)

## [US National Virtual Observatory](#)

The National **Virtual Observatory**: Tools and Techniques for Astronomical Research, ASP Vol. 382), is available for pre-order at the ASP website. ...

[www.us-vo.org/](http://www.us-vo.org/) - 16k - [Cached](#) - [Similar pages](#)

## [International Virtual Observatory Alliance](#)

IVOA logo International **Virtual Observatory** Alliance. About IVOA · Members · Contacts · IVOA Executive · Working Groups · Documents and Standards ...

[www.ivoa.net/](http://www.ivoa.net/) - 12k - [Cached](#) - [Similar pages](#)

## [The National Virtual Observatory \(NVO\)](#)

Welcome to the education site of the National **Virtual Observatory**! ... Developed in collaboration with the International **Virtual Observatory** Alliance. ...

[www.virtualobservatory.org/](http://www.virtualobservatory.org/) - 20k - [Cached](#) - [Similar pages](#)

## [European Virtual Observatory](#)

The EURO-VO project aims at deploying an operational **Virtual Observatory** (VO) in Europe. Its objectives are technology take-up and VO compliant resource ...

[www.euro-vo.org/](http://www.euro-vo.org/) - 22k - [Cached](#) - [Similar pages](#)

## [Virtual Observatory India](#)

**Virtual Observatory** India.

[vo.iucaa.ernet.in/](http://vo.iucaa.ernet.in/) - 7k - [Cached](#) - [Similar pages](#)

## [Australian Virtual Observatory](#)

One concept that is being developed by a number of countries is that of the **virtual observatory** (VO). These pages are intended to explain the concept, ...

[www.atnf.csiro.au/projects/avo/](http://www.atnf.csiro.au/projects/avo/) - 11k - [Cached](#) - [Similar pages](#)

## [Theory in a Virtual Observatory](#)

Theory in a **Virtual Observatory**. This page is a growing list of links to different type of theory data that various researchers have made available. ...

[bima.astro.umd.edu/nemo/tvo/](http://bima.astro.umd.edu/nemo/tvo/) - 7k - [Cached](#) - [Similar pages](#)

## [The <I>Virtual Observatory </I> CDS page](#)

The aim of the **Virtual Observatory** (VO) is to allow astronomers immediate access to the data archives, reference surveys, data bases and information ...

[cdsweb.u-strasbg.fr/avo.htx](http://cdsweb.u-strasbg.fr/avo.htx) - 9k - [Cached](#) - [Similar pages](#)

## [Virtual Observatory - Wikipedia, the free encyclopedia](#)

A **virtual observatory** is a collection of interoperating data archives and software tools which utilize the internet to form a scientific research ...

[en.wikipedia.org/wiki/Virtual\\_Observatory](http://en.wikipedia.org/wiki/Virtual_Observatory) - 19k - [Cached](#) - [Similar pages](#)

# ...and “the BIRN”?

**BIRN** Biomedical Informatics Research Network  
*fostering a new biomedical collaborative culture and infrastructure*

Search  
  
● BIRN ● WWW

[Home](#) [Research](#) [Tools](#) [Data](#) [Contact Us](#) [Help](#) [BIRN Portal](#)

**BIRN Collaborations**  
[BIRN Multi-site Collaborations](#)  
[Information on NIH Program Announcements](#)  
[News and Events](#)  
[Comments and Suggestions](#)

**BIRN Research**  
[BIRN Coordinating Center](#)  
[Morphometry BIRN](#)  
[Function BIRN](#)  
[Mouse BIRN](#)

**BIRN Tools**  
[Browse Our Tools](#)  
[Use Our Tools](#)  
[Share Your Tools](#)

**BIRN Data Repository**  
[View Data Currently Available](#)  
[Preview Data Coming Soon](#)  
[Data Use](#)  
[Share Your Data](#)

[About Us](#) [Copyright](#) [Privacy](#) [Site Map](#) [Contact](#) [User Survey](#)

BIRN is supported by NIH Grants to the BIRN Coordinating Center (U24-RR019701), Function BIRN (U24-RR021992), Morphometry BIRN (U24-RR021382), and Mouse BIRN (U24-RR021760).



(see Jeffrey Grethe’s talk in this Symposium)

# BIRN Organization



# What is “the” Virtual Observatory?

[Advanced Search](#)  
[Preferences](#)

## Web

### [SkyView Virtual Observatory](#)

**Virtual observatory** on the net generating images of any part of the sky at wavelengths in all regimes from radio to gamma-ray.

[skyview.gsfc.nasa.gov/](#) - 10k - [Cached](#) - [Similar pages](#)

### [US National Virtual Observatory](#)

The National **Virtual Observatory**: Tools and Techniques for Astronomical Research, ASP Vol. 382), is available for pre-order at the ASP website. ...

[www.us-vo.org/](#) - 16k - [Cached](#) - [Similar pages](#)

### [International Virtual Observatory Alliance](#)

IVOA logo International **Virtual Observatory** Alliance. About IVOA · Members · Contacts · IVOA Executive · Working Groups · Documents and Standards ...

[www.ivoa.net/](#) - 12k - [Cached](#) - [Similar pages](#)

### [The National Virtual Observatory \(NVO\)](#)

Welcome to the education site of the National **Virtual Observatory**! ... Developed in collaboration with the International **Virtual Observatory** Alliance. ...

[www.virtualobservatory.org/](#) - 20k - [Cached](#) - [Similar pages](#)

### [European Virtual Observatory](#)

The EURO-VO project aims at deploying an operational **Virtual Observatory** (VO) in Europe. Its objectives are technology take-up and VO compliant resource ...

[www.euro-vo.org/](#) - 22k - [Cached](#) - [Similar pages](#)

### [Virtual Observatory India](#)

**Virtual Observatory** India.

[vo.iucaa.ernet.in/](#) - 7k - [Cached](#) - [Similar pages](#)

### [Australian Virtual Observatory](#)

One concept that is being developed by a number of countries is that of the **virtual observatory** (VO). These pages are intended to explain the concept, ...

[www.atnf.csiro.au/projects/avo/](#) - 11k - [Cached](#) - [Similar pages](#)

### [Theory in a Virtual Observatory](#)

Theory in a **Virtual Observatory**. This page is a growing list of links to different type of theory data that various researchers have made available. ...

[bima.astro.umd.edu/nemo/tvo/](#) - 7k - [Cached](#) - [Similar pages](#)

### [The <I>Virtual Observatory </I> CDS page](#)

The aim of the **Virtual Observatory** (VO) is to allow astronomers immediate access to the data archives, reference surveys, data bases and information ...

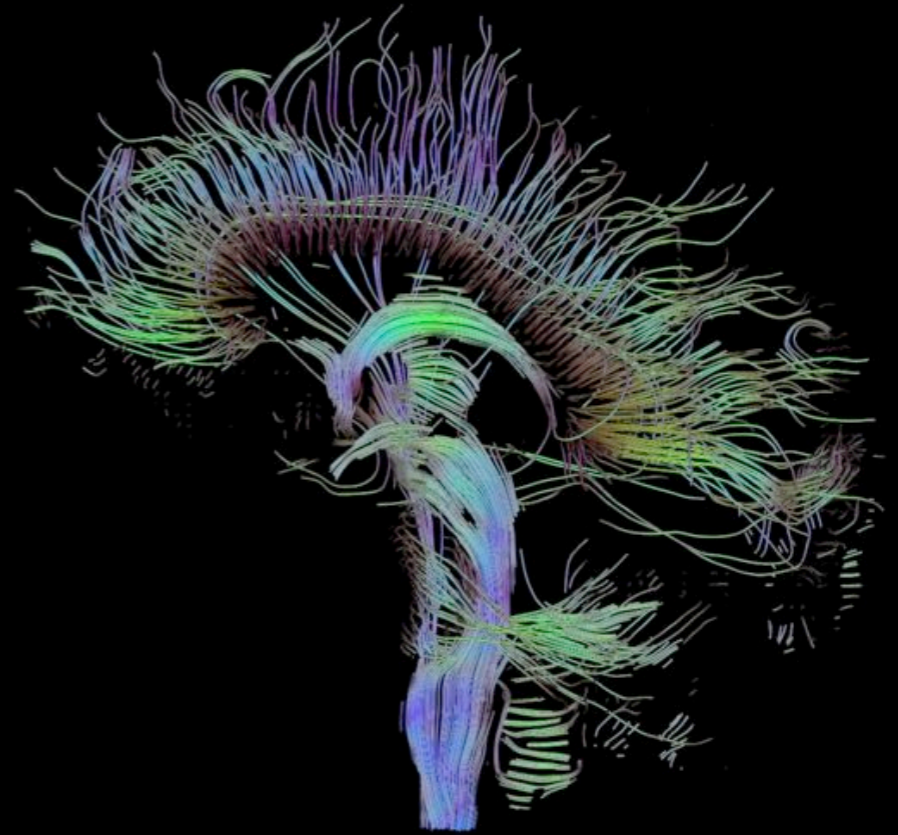
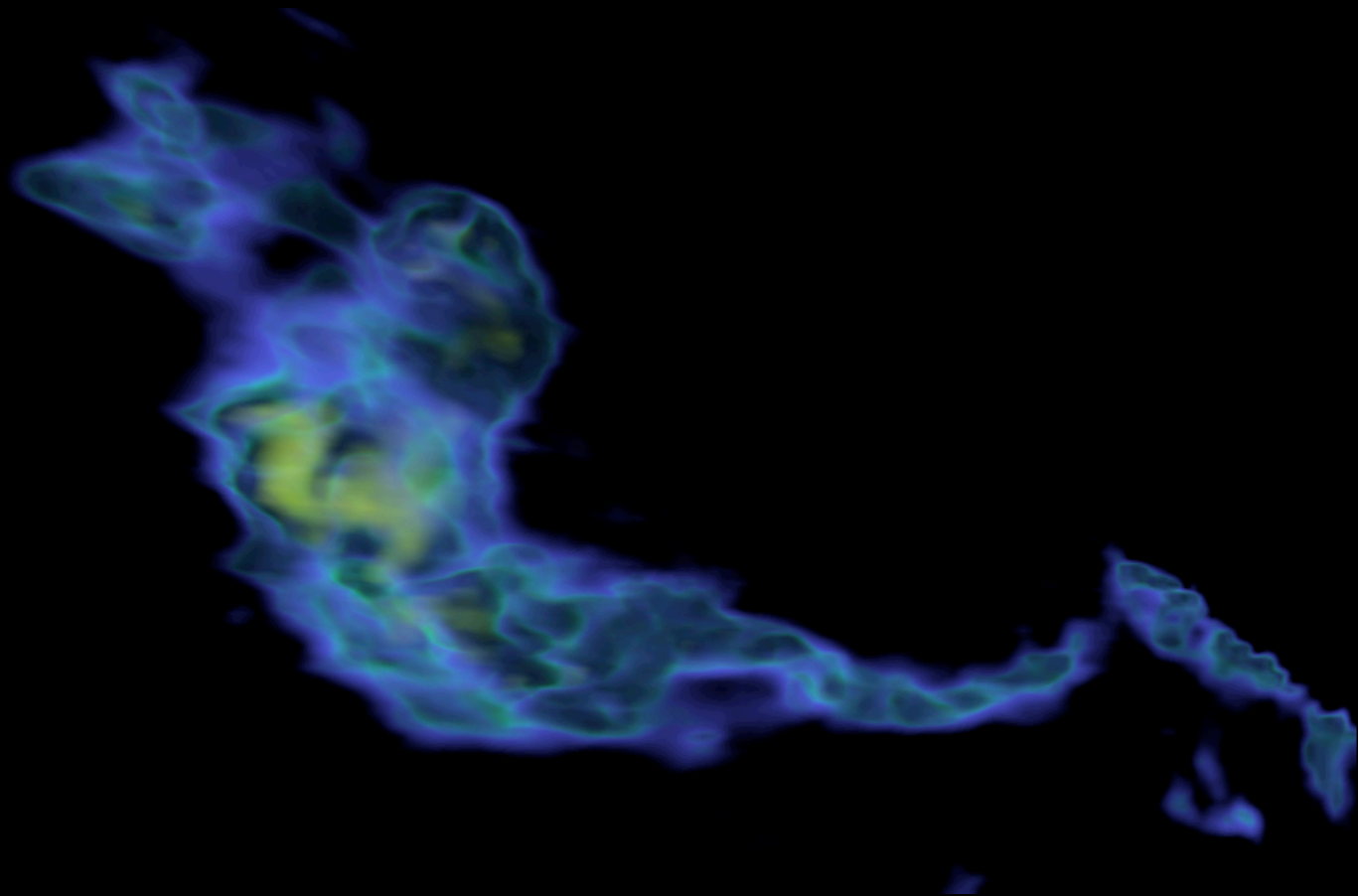
[cdsweb.u-strasbg.fr/avo.htx](#) - 9k - [Cached](#) - [Similar pages](#)

### [Virtual Observatory - Wikipedia, the free encyclopedia](#)

A **virtual observatory** is a collection of interoperating data archives and software tools which utilize the internet to form a scientific research ...

[en.wikipedia.org/wiki/Virtual\\_Observatory](#) - 19k - [Cached](#) - [Similar pages](#)

# Data Intensive Science in Astronomy & Medicine



Alyssa A. Goodman  
Initiative in Innovative Computing @ Harvard  
and



Harvard-Smithsonian Center for Astrophysics