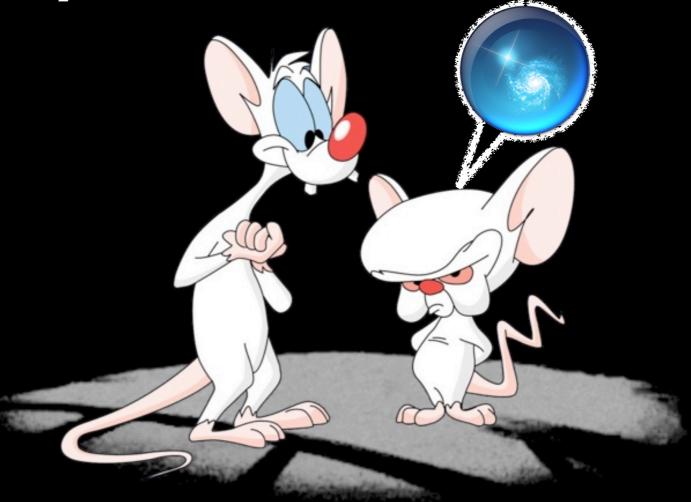
# Seamless Astronomy Enabled by WWT



## Alyssa A. Goodman Harvard-Smithsonian Center for Astrophysics

## The Slide to Rule them All...



#### Collaborators:

Alberto Accomazzi, Douglas Burke, Raffaele D'Abrusco, Rahul Davé, Christopher Erdmann, Pepi Fabbiano, Alyssa Goodman, Jay Luker, Gus Muench, Michael Kurtz & Alberto Pepe (Harvard-Smithsonian CfA); Eli Bressert (U. Exeter); Tim Clark (Massachusetts General Hospital/Harvard Medical School); Mercé Crosas (Harvard Institute for Quantitative Social Science; Chris Borgman (UCLA); Jonathan Fay & Curtis Wong (Microsoft Research)

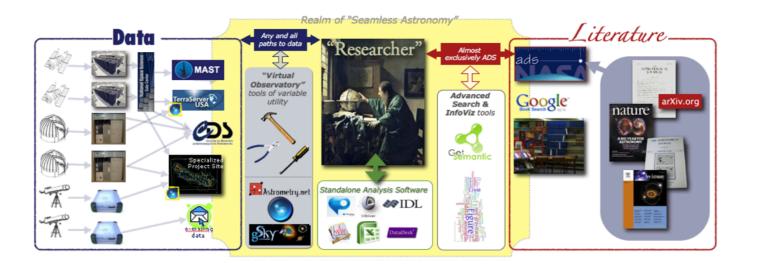






ABOUT PEOPLE PROJECTS PUBLICATIONS PRESENTATIONS SOFTWARE CFA DATA (BETA)

#### 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.

#### Sponsors of Seamless Astronomy include NASA, NSF and Microsoft Research.

Contact us. For inquiries or questions, please email Sarah Block at sblock@cfa.harvard.edu. Alternatively you can contact or visit us at: SEAMLESS ASTRONOMY TEAM HARVARD-SMITHSONIAN CENTER FOR ASTROPHYSICS 60 GARDEN STREET, MS 42 CAMBRIDGE, MA 02138 Events Twitter Links

#### Announcements

#### Latest news

augustmuench: Farmers' Almanac foresees a rough winter ahead & dowser finds huge body of dirty water under the Longfellow bridge http://t.co/DwL2mMT

#### albertoconti: RT

@james\_s\_bullock: Why doesn't the History Channel just change its name to the Bigfoot Lover's Pseudoscience Channel?http://j.mp/rlKp4C

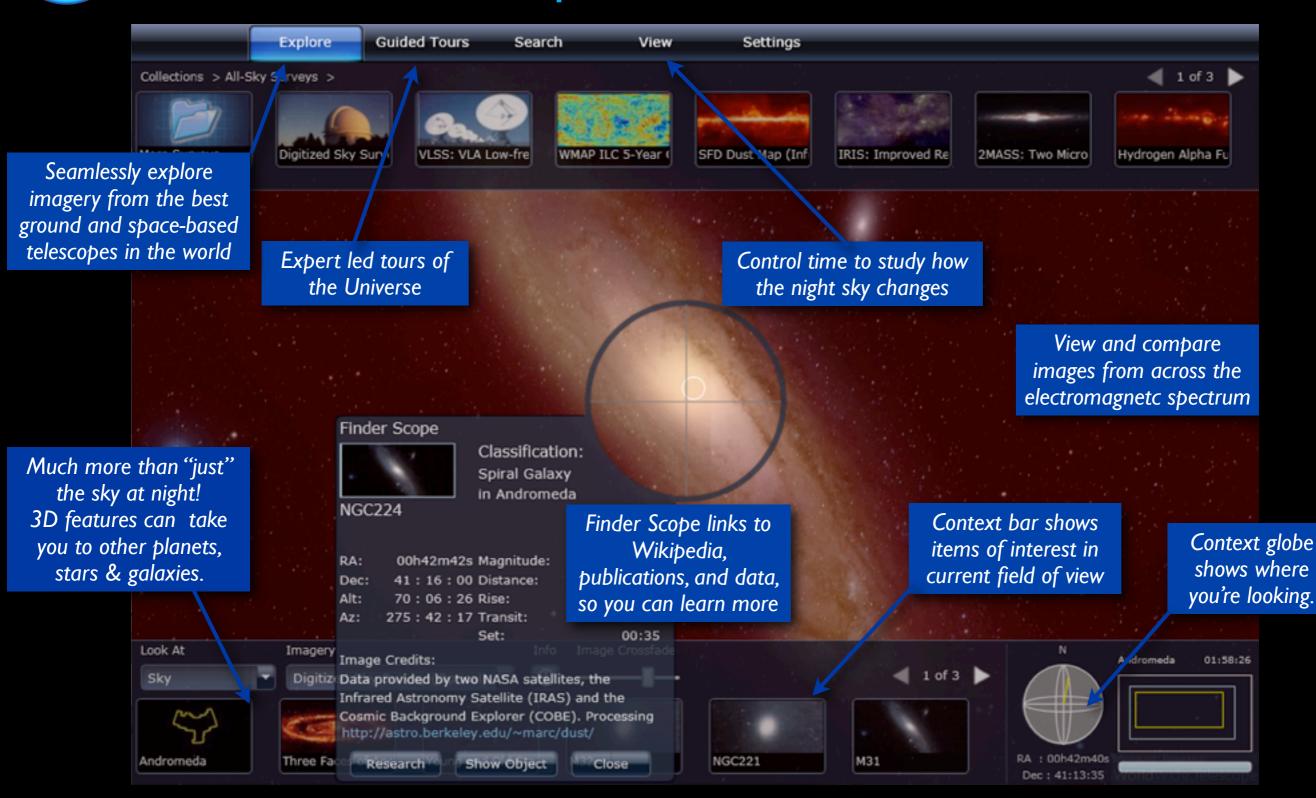
albertoconti: RT @johnmaeda: "Not everything knowable can be articulated in propositional form." http://t.co/ZYD43ER

albertoconti: RT @sarahkendrew: looks awesome! >> @astrobetter: New Post: iObserve: The Astronomical Observing App We've Been Waiting For http:/ ...

augustmuench: my post `on open science and anonymous peer review`

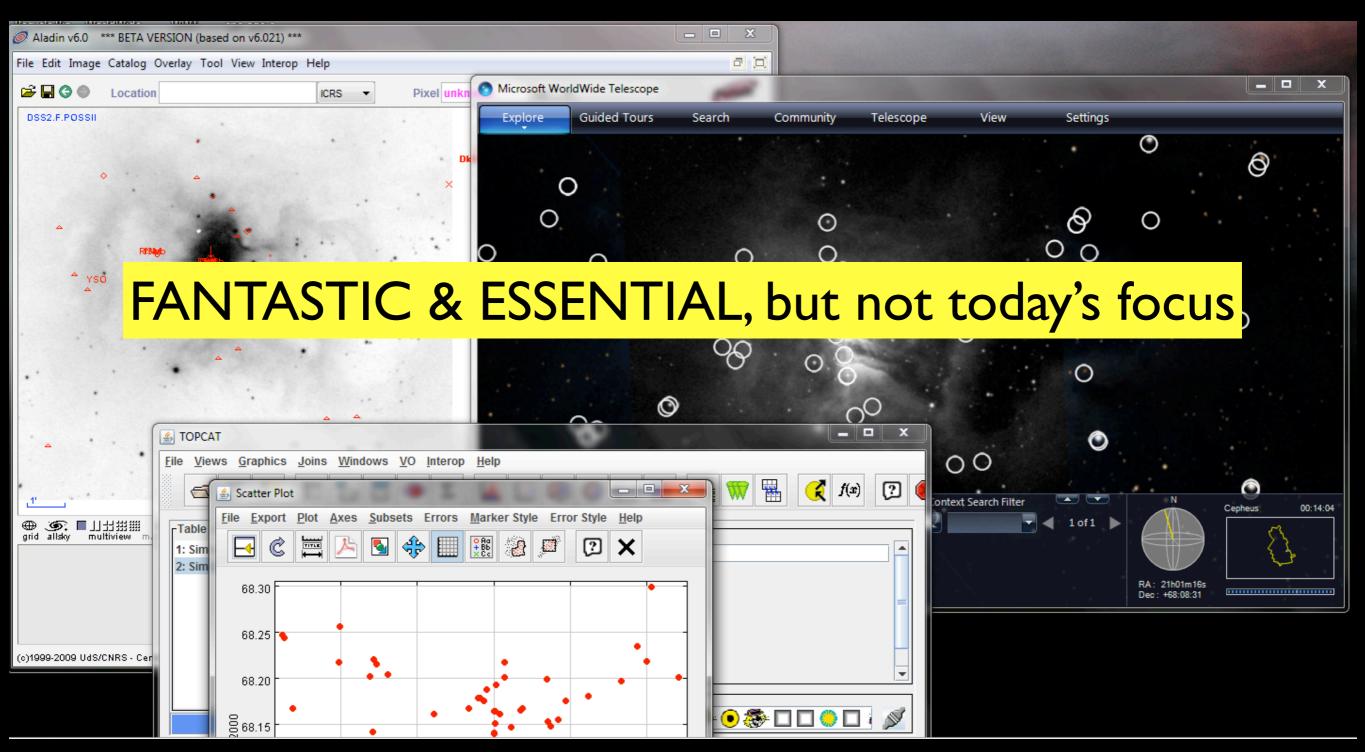
## Microsoft<sup>®</sup> Research WorldWide Telescope

## Experience WWT at worldwidetelescope.org



SAMP

(Simple Application Messaging Protocol)



link to 12/2010 IVOA recommendation

Microsoft<sup>®</sup> Research WorldWide Telescope

Ambassadors Program



## FANTASTIC & ESSENTIAL, but not today's focus



## wwtambassadors.org

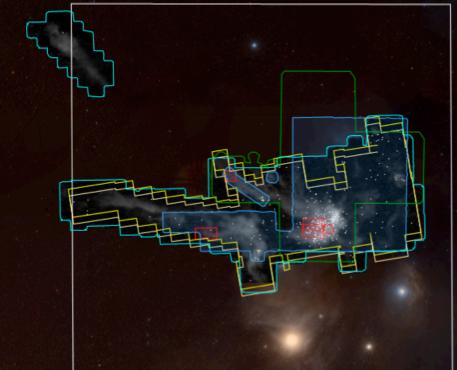


worldwidetelescope.org labs.adsabs.harvard.edu/ui/

# (My) Research



## Perseus



## Ophiuchus

orldWide Telescope

Serpens

# C C P L E T E

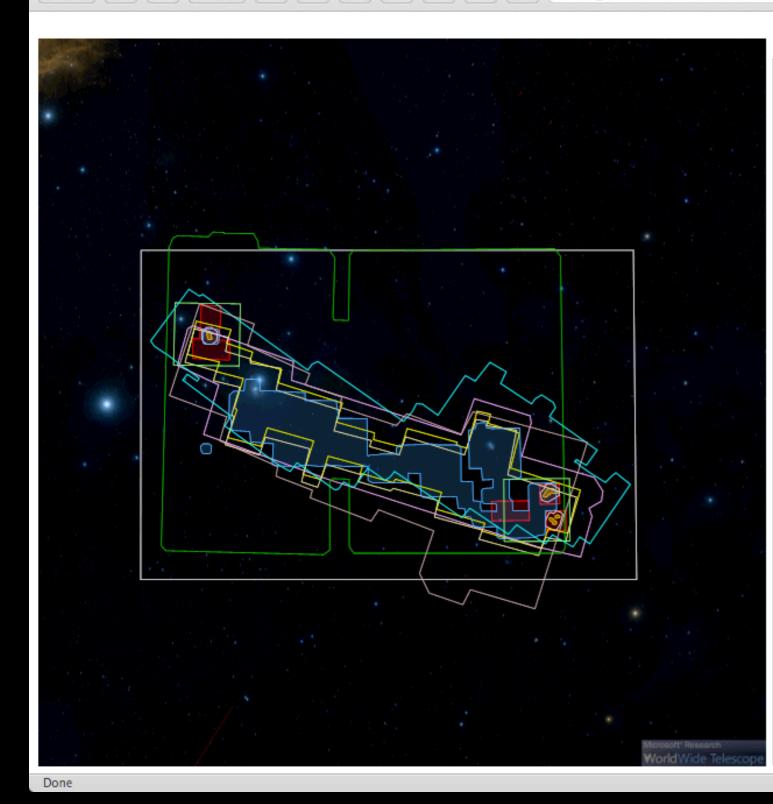
The **CO**ordinated Molecular Probe Line Extinction Thermal Emission Survey of Star-Forming Regions

## www.cfa.harvard.edu/COMPLETE tinyurl.com/completepapers

1P

+

0



 $\Diamond$ 

Î

#### **COMPLETE Data Available**

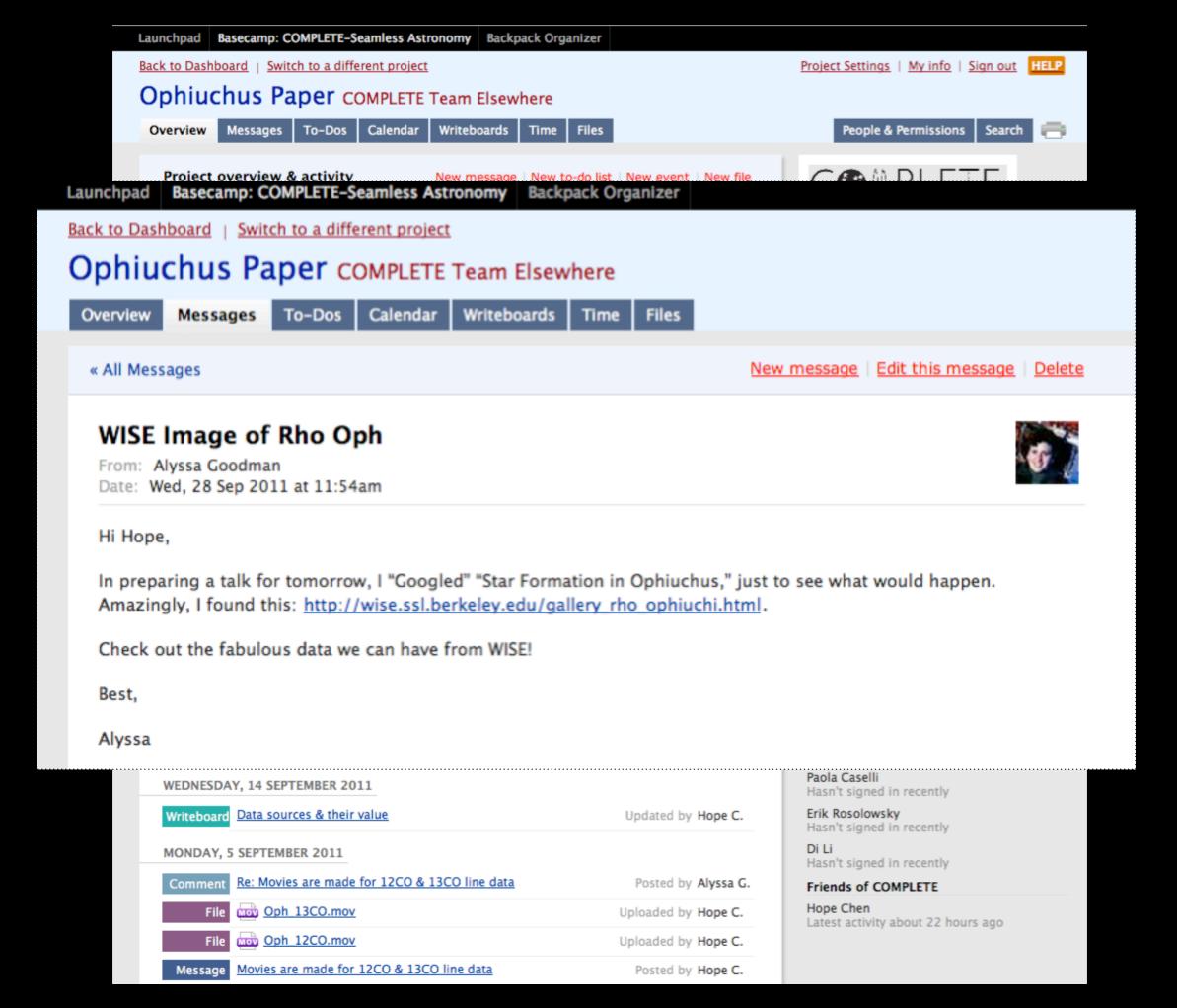
Cen	ter on Perseus Ce	nter on O	phichus	Center on Serper	ns				
Full-Cloud Data (Phase I, All Data Available)									
Dataset		Show	Perseus	Ophiuchus	Serpens	Link			
GBT: HI Data Cube	<del>}</del>	$\sim$	۷	٧	Ø	Data			
IRAS: Av/Temp Ma	aps	$\sim$	۷	٧	۷	Data			
FCRAO: 12CO		$\mathbf{V}$	<b>⊻</b>	<u>v</u>	۷	Data			
FCRAO: 13CO			٧	<u>⊻</u>	۷	Data			
JCMT: 850 micron	5	$\sim$	٧	<u>⊻</u>	Ø	Data			
Spitzer c2d: IRAC	1,3 (3.6,5.8 µm)	$\sim$	۷	٧	۷	Data			
Spitzer c2d: IRAC	2,4 (4.5,8 μm)	$\mathbf{V}$	۷	٧	<b>⊻</b>	Data			
CSO/Bolocam: 1.2	-mm	$\sim$	⊻	Ø	Ø	<u>Data</u>			
Spitzer MIPS: Deri	ved Dust Map	$\mathbf{V}$	٧	Ø	Ø	Data			
Target	ed Regions (Pha	se II, So	me Data	Not Yet Availa	able)				
CTIO/Calar Alto: N	llR (J,H,Ks)	$\mathbf{N}$	۷	٧	Ø	Data			
IRAM 30-m: N2H+	and C18O	$\mathbf{V}$	۷	Ø	Ø	Data			
IRAM 30-m: 1.1-n	nm continuum	N	۷	Ø	Ø	Data			
Megacam/MMT: r	,i,z images	N	۷	Ø	Ø	Data			
Catalogs & Pointed Surveys									
NH3 Pointed Surv	ey		۷	Ø	Ø	Data			
YSO Candidate list	t (c2d)		۷	٧	٧	<u>Data</u>			

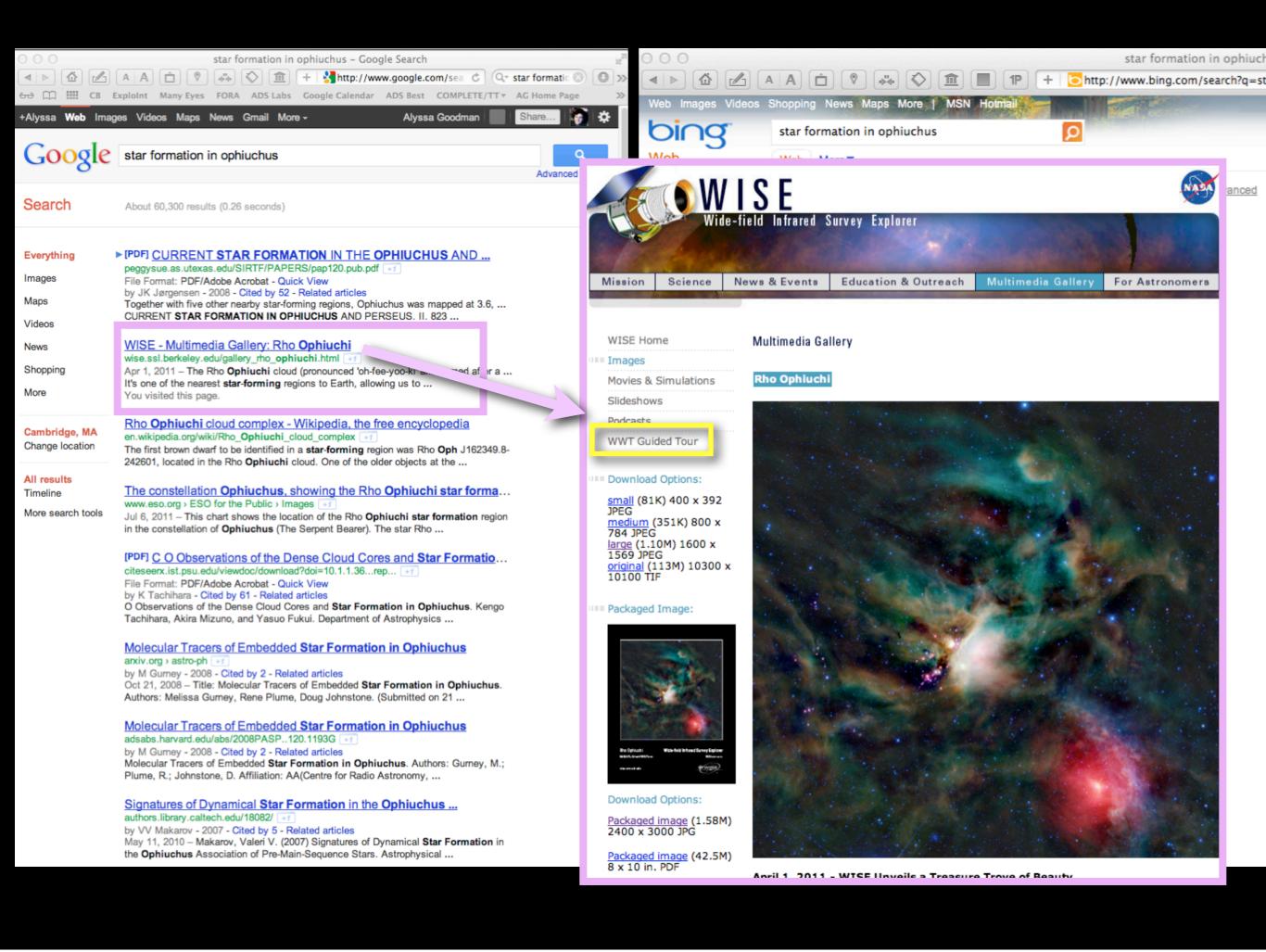
http://www.worldwidetelescope.org/COMPLETE/WWTCoverageTool.htm

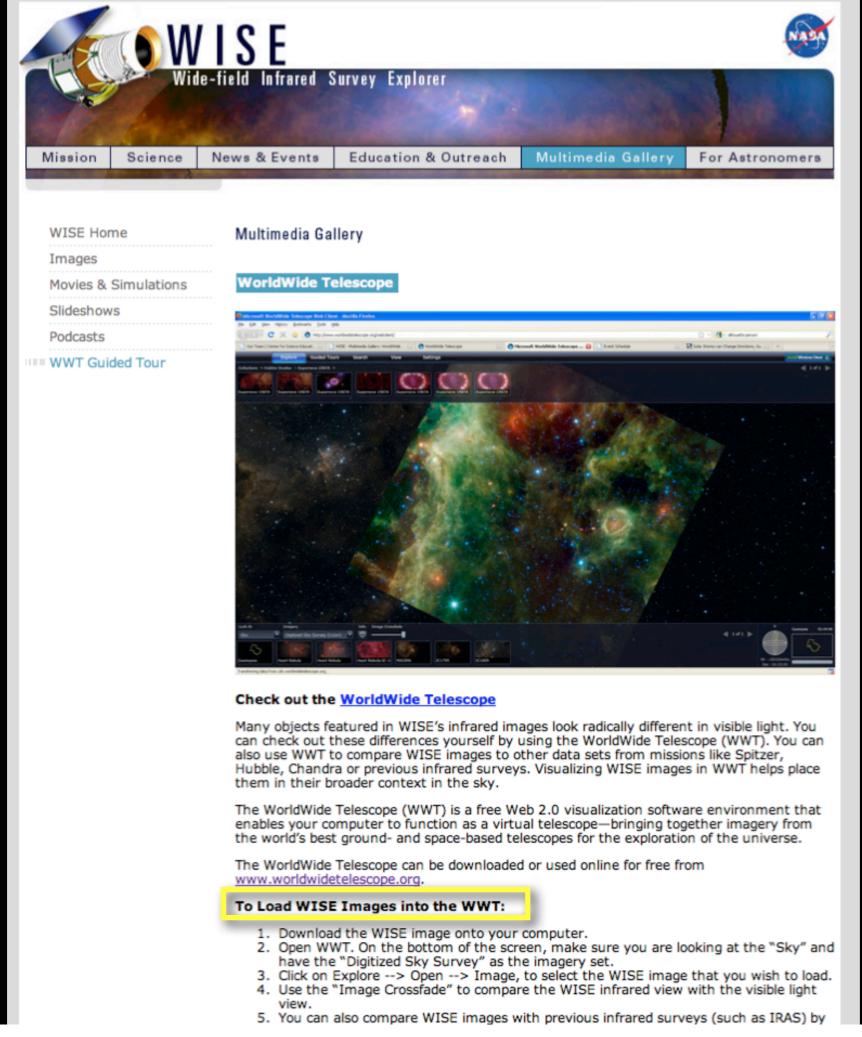
# A True Story



## Hope Chen Brand-new Harvard Grad Student Project: "COMPLETE" Ophiuchus



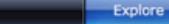






Finder Scope					
Classification:					
Star in Ophiuchus					
Rho Ophiuchi; 5 Ophiuchi; HR6112; SAO184382; HD147933; DM					
RA: 16h25m35s Magnitude: 5.0	02				
Dec: -23 : 26 : 50 Distance: n/	n/a				
Alt:         -02 : 13 : 24 Rise:         12:0           Az:         120 : 16 : 22 Transit:         16:3					
Set: 21:1	11				
Image Credits: Copyright DSS Consortium					
http://gsss.stsci.edu/Acknowledgements/DataCo	Co				
Research Show Object Close					





Collections >

Constellations

Guided Tours

All-Sky Surveys

**\_**//••

Solar System (Sky

Settings

Chandra Studies



Hubble Studies











Finder Scope



Search

Classification: Star

View

Spitzer Studies

Rho Ophiuchi; 5 Ophiuchi; HR6112; SAO184382; HD147933; DM

RA:	16h		5s	Magnitude:	5.0
Dec:	-23 :	26:5	50	Distance:	n/
Alt:	01:	22:2	28	Rise:	12:0
Az:	235 :	51:1	12	Transit:	16:3
				Set	21+1

Image Credits:

Copyright DSS Consortium

http://gsss.stsci.edu/Acknowledgements/DataCo

		Research Show Object Close Name:Rho Ophiuchi;5 Ophiuchi;HR		SIMBA
O. over	rlavs	Information Imagery Virtual Observatory Searches		Look up on SIMBAD Look up on SEDS Look up on Wikipedia
Imagery Digitized Sky Su		Set as Foreground Imagery Set as Background Imagery		Look up publications o Look up on NED Look up on SDSS
1C4603	1C4604	Properties Copy Shortcut Share on Facebook	01845	Omega Ophiuchi;9

2.1						
	Look up on SIMBAD					
	Look up on SEDS					
	Look up on Wikiped	lia 🔥 🔥				
	Look up publication	is on ADS $A$	DS	N	Ophiuchus	06:25:36
	Look up on NED		🚽 1 of 2 🕨			
	Look up on SDSS					
4018 <b>45</b> 4	Omega Ophiuchi;9	Psi Ophiuchi;4 Oph	Rho Ophiuchi;5 Op	RA : 16h25m17s	-	_

#### Tuesday, December 6, 2011

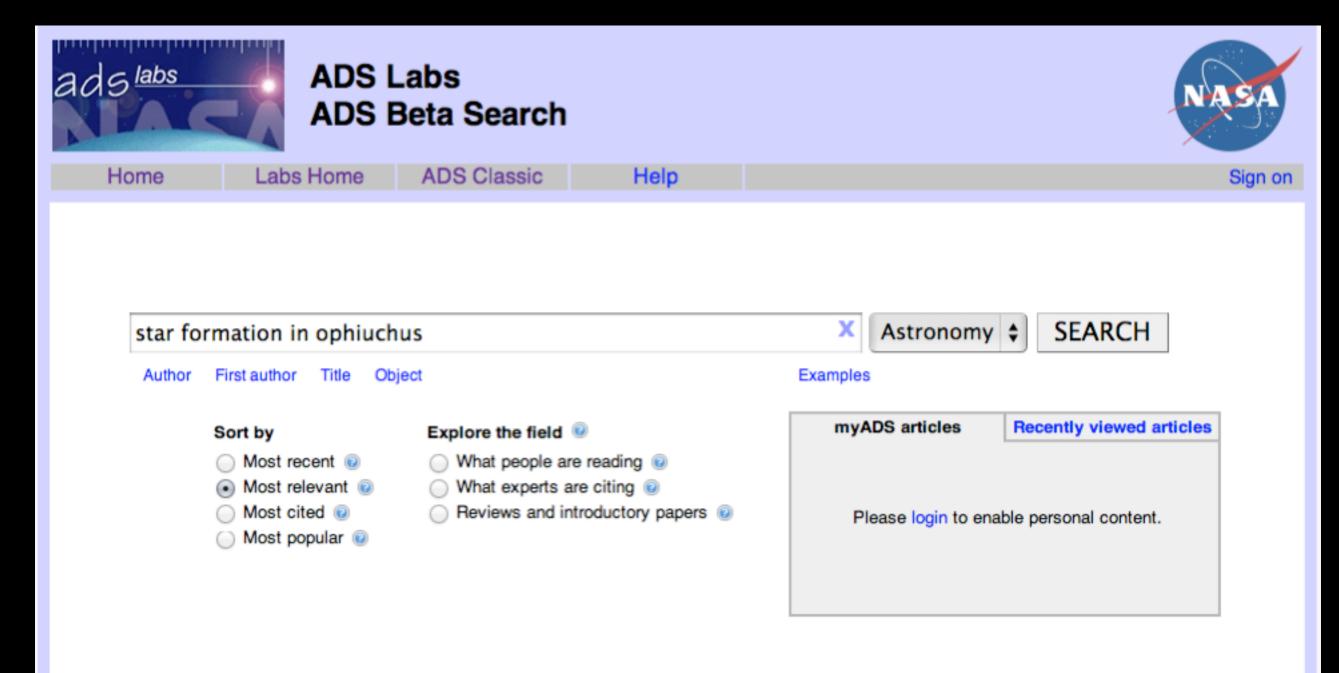
Look At

Ophiuchus

## Choosing ADS link gives...

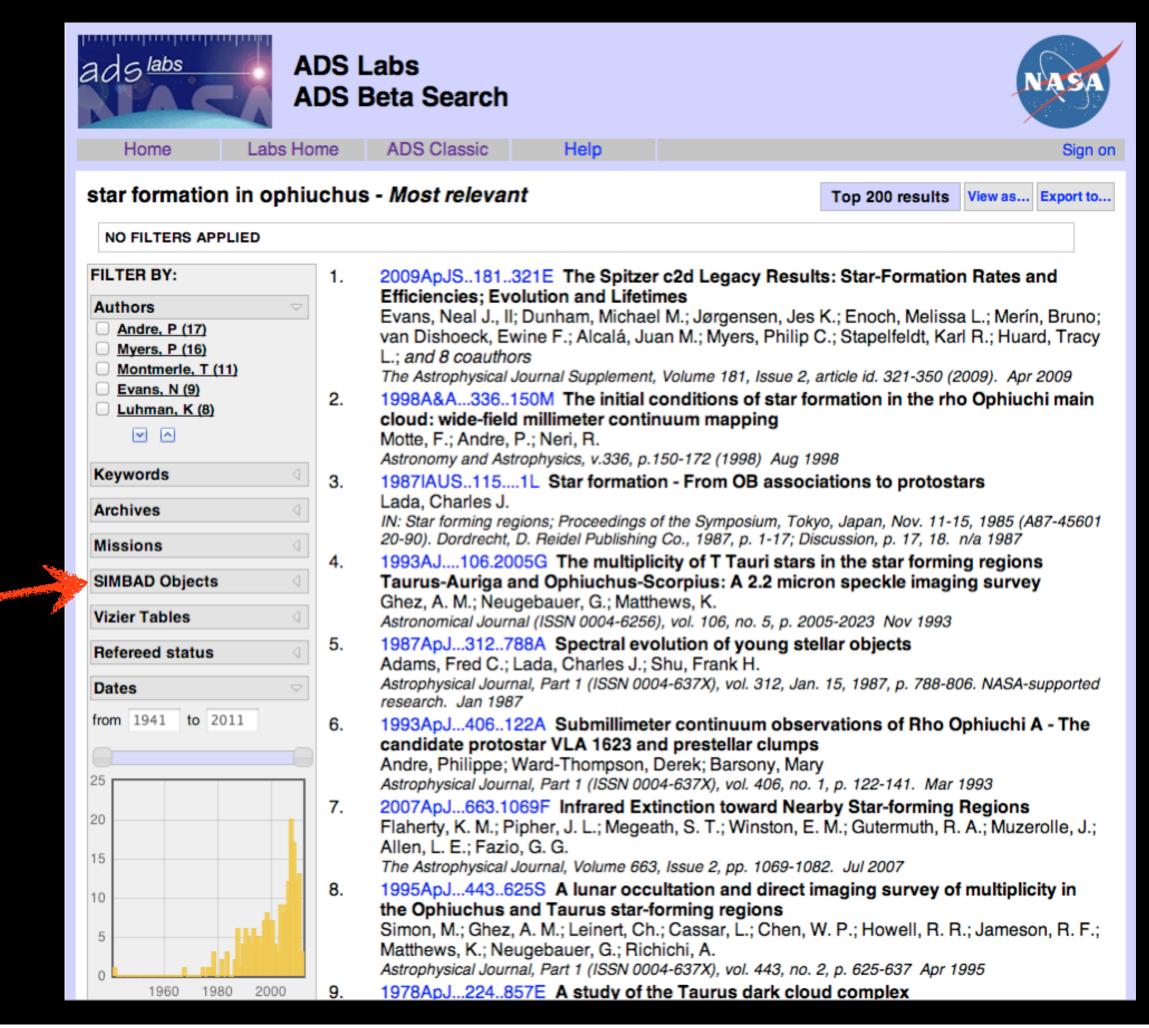
		201									Sign on
SAO/NASA Astrophysics Data System (ADS) Query Results from the Astronomy Database								Go to bottom of page			
Re	etrieved <b>200</b> abstracts, starting with number <b>1</b> .	Total numb	er selected: 439.								Sort options \$
#	Bibcode Authors	Score Title	Date		st of Lin cess Co	_	leip				
1	<ul> <li><u>2011ApJ738115D</u></li> <li>del Valle, María Victoria;</li> <li>Romero, Gustavo E.; Luque-</li> </ul>	1.000	09/2011 Iri Stars Gamma-I	A	<u>e</u> e	LΧ		R	<u>5</u>	<u>U</u>	
	Escamilla, Pedro Luis; Martí, Josep; Ramón Sánchez-Sutil, Juan										
2	2011ApJ73796G	1.000	08/2011	Α	E E	LX		<u>R</u> <u>C</u>	<u>s</u>	U	
	Goldsmith, Paul F.; Liseau, René; Bell, Tom A.; Black, John H.; Chen, Jo-Hsin; Hollenbach, David; Kaufman, Michael J.; Li, Di; Lis, Dariusz C.; Melnick, Gary; and 25 coauthors	Herschel	Measurements of	f Mo	lecular	Oxyge	n in Orio	on			
3	2011ApJ73465J	1.000	06/2011	Α	ΕE	LX	D	<u>R</u> <u>C</u>	<u>s</u>	U	
	Jenkins, Edward B.; Tripp, Todd M.		bution of Therma ture Excitations	l Pre	essures i	in the I	Diffuse,	Cold Neutr	al Med	lium of Our Galaxy. I	I. An Expanded Survey of Interstellar C I
4	2011AJ141201M	1.000	06/2011	Α	ΕE	LΧ	D	<u>R</u>	<u>s</u>	U	
	McCleary, J. E.; Wolk, S. J.	A Survey	of High-contrast S	stella	ar Flares	Obser	ved by	Chandra			
5	2011ApJ7321011	1.000	05/2011	A	ΕE	LX		<u>R</u> <u>C</u>	<u>s</u>	<u>U</u>	
	Ikeda, Norio; Kitamura, Yoshimi	Similarity	Between the C18	() O	= 1-0) 0	Core M	ass Fund	ction and t	he Initia	al Mass Function (IM	F) in the S140 Region
6	2011AJ141165A	1.000	05/2011	Α	ΕE	L		R	<u>s</u>	U	
	Abt, Helmut A.	The Age of	of the Local Inters	tella	r Bubbl	е					

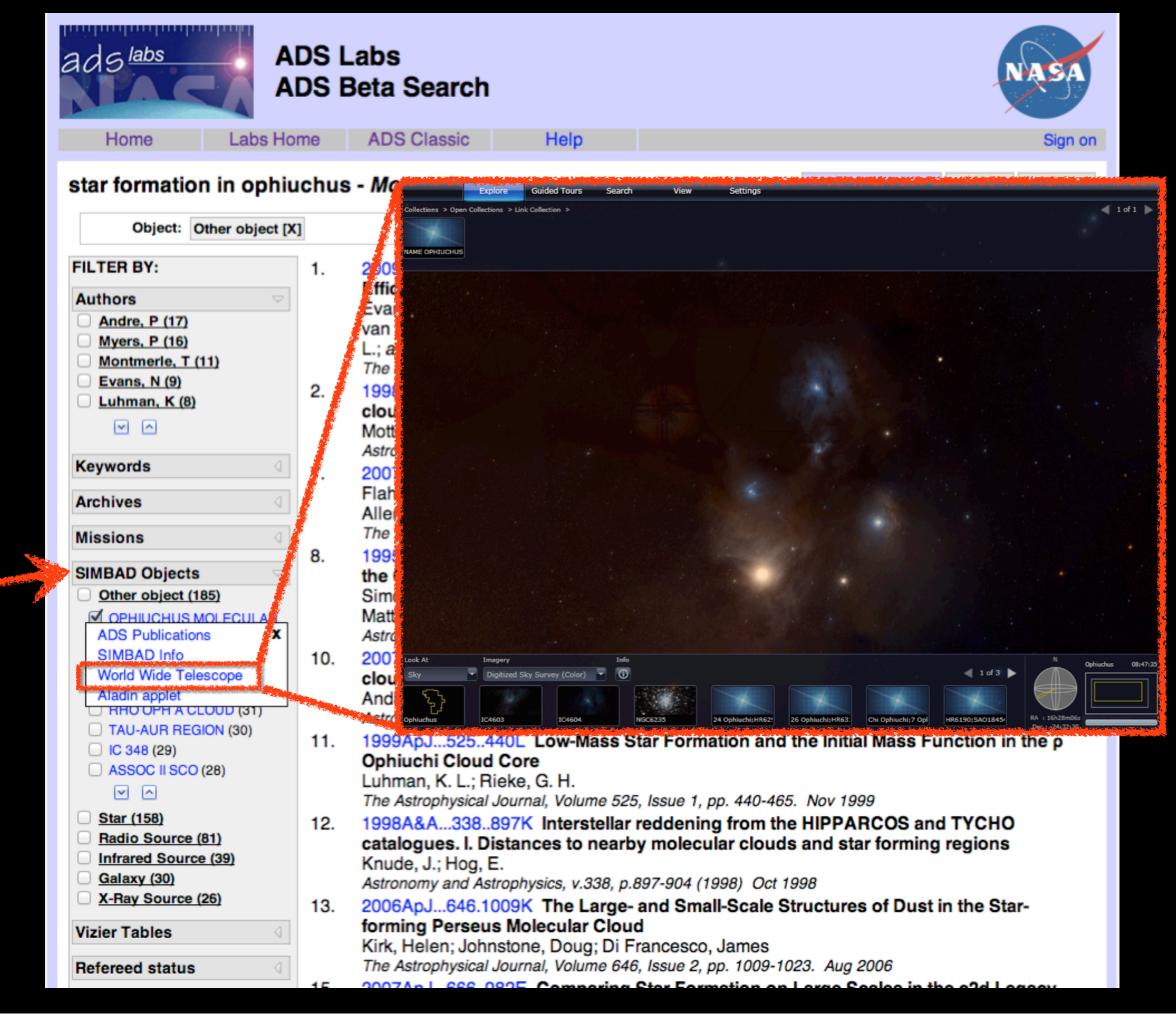
## Starting with ADS Labs gives...





The ADS is operated by the Smithsonian Astrophysical Observatory under NASA Grant NNX09AB39G FOLLOW US ON EWITTER Like 671 Contact: ads at cfa.harvard.edu or through the feedback form.



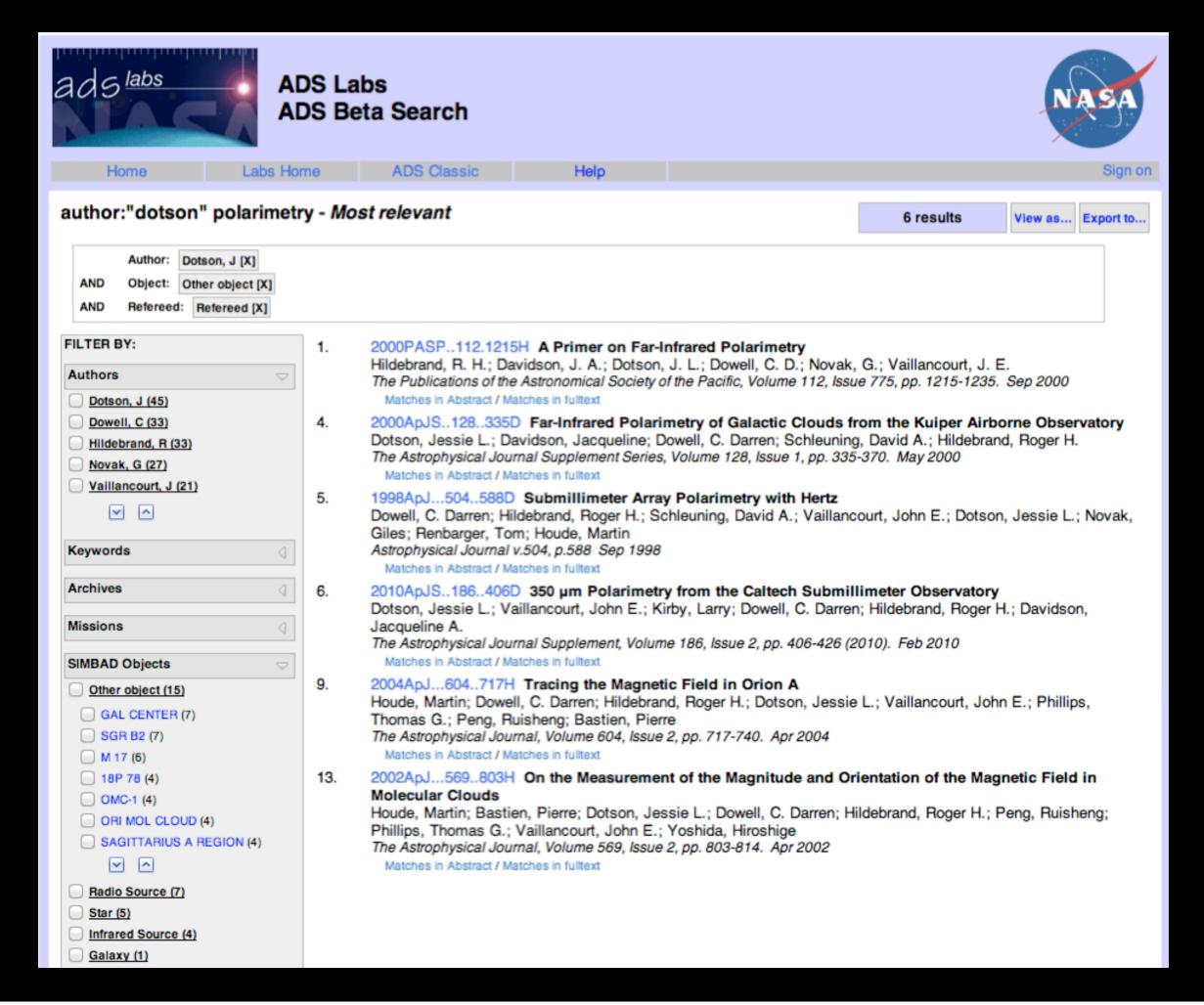


# Also yesterday...



## Chris Beaumont Hawaii-Harvard Grad Student Proposal: M17 Polarimetry





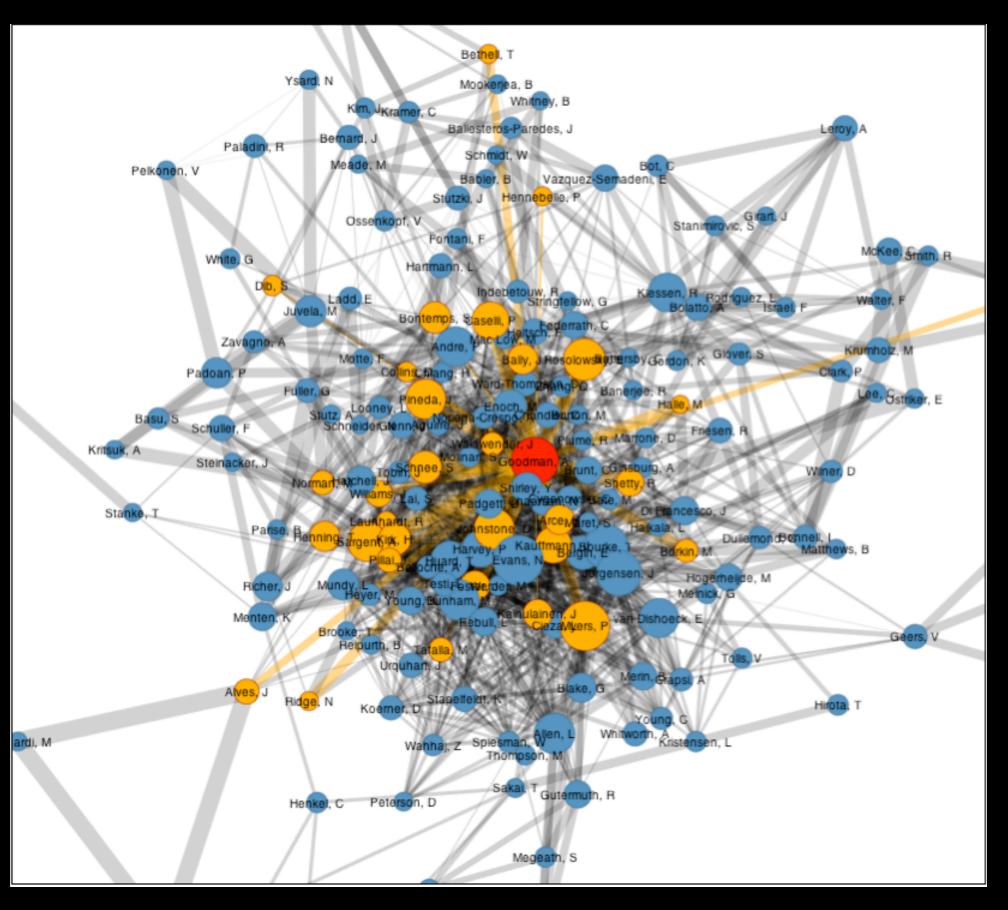
# back to... COMPLETE

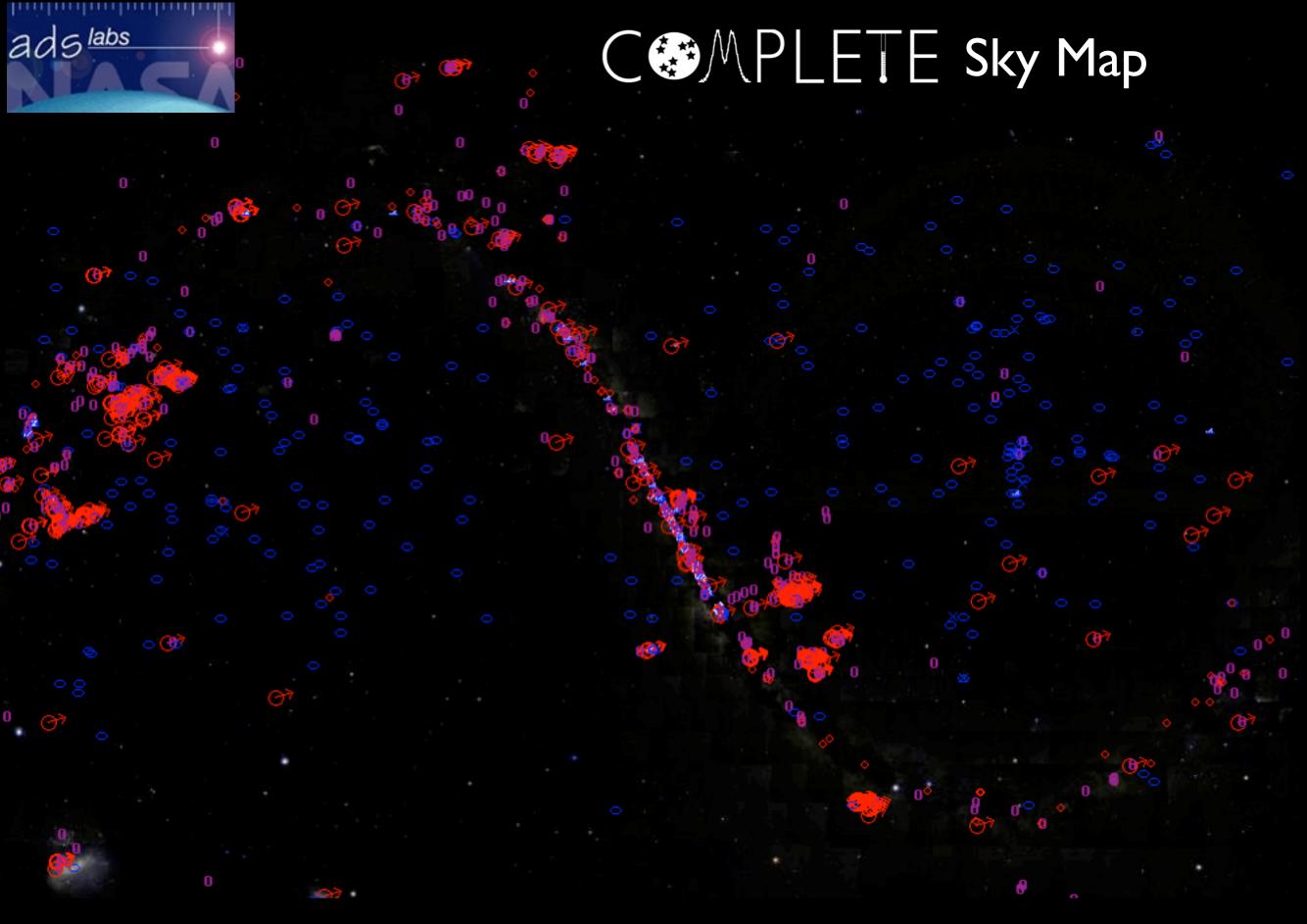


adio	DS L DS B	abs leta Search	1			NASA			
Home Labs Ho	me	ADS Classic	Help			Sign on			
<b>COMPLETE Papers - P</b> NO FILTERS APPLIED	Private	Library			36 results	View as Export to Author network			
FILTER BY:	1.	2011arXiv1109.	3368A A Bubbling	Nearby Molecular C	loud: COMPLE	Paper network			
Authors <u>Goodman, A (25)</u> <u>Pineda, J (12)</u> <u>Johnstone, D (10)</u> <u>Kirk, H (10)</u>	2.	Perseus Arce, Hector G.; Christopher N. eprint arXiv:1109.3 2011ApJ739L	Borkin, Michelle A. 3368 Sep 2011	; Goodman, Alyssa A.; ery Large Array Obse	; Pineda, Jaime	Sky Map			
□ Foster, J (9) ☑ △		Pineda, Jaime E Corder, Stuartt	.; Goodman, Alyssa	a A.; Arce, Héctor G.; C me 739, Issue 1, article i					
Keywords	3.	Observations							
Missions d		Goodman, Alyss Computational Sta Volume 270, p. 51	on, IAU Symposium,						
SIMBAD Objects	4.	2010ApJ723.1019H The Star Formation Rate and Gas Surface Density Relation in the Milky Way: Implications for Extragalactic Studies							
Vizier Tables		Heiderman, Ama	anda; Evans, Neal J	I., II; Allen, Lori E.; Hua Issue 2, article id. 1019					
Dates ~	5.	2010ApJ723457K The Dynamics of Dense Cores in the Perseus Molecular Cloud. II. The Relationship Between Dense Cores and the Cloud Kirk, Helen; Pineda, Jaime E.; Johnstone, Doug; Goodman, Alyssa							
from 2004 to 2011				stone, Doug; Goodma Issue 1, pp. 457-475 (2					
	6.		433K The Mass-si	ze Relation from Clo		Solar			
8		The Astrophysical	Journal, Volume 716	Myers, P. C.; Goodma Issue 1, pp. 433-445 (2	010). Jun 2010				
6	7.	Arce, Héctor G.; W.	Borkin, Michelle A.	ETE Survey of Outfle Goodman, Alyssa A.	; Pineda, Jaime I	E.; Halle, <mark>Michael</mark>			
4	8.	2010ApJ712.1 Structure in Mo Kauffmann, J.; P	137K The Mass-S lecular Clouds Villai, T.; Shetty, R.;	Issue 2, pp. 1170-1190 ize Relation from Clo Myers, P. C.; Goodma Issue 2, pp. 1137-1146	n, A. A.	. A New Probe of			
2004 2006 2008 2010 2012	9.	Size Relationsh	<b>hips</b> Collins, David C.; Ka	of Projection on Deriv					

## COMPLETE Citation Network (from ADS Labs)

many thanks to A. Accomazzi, R. Davé, M. Kurtz, G. Di Milia, A. **Pepe** 



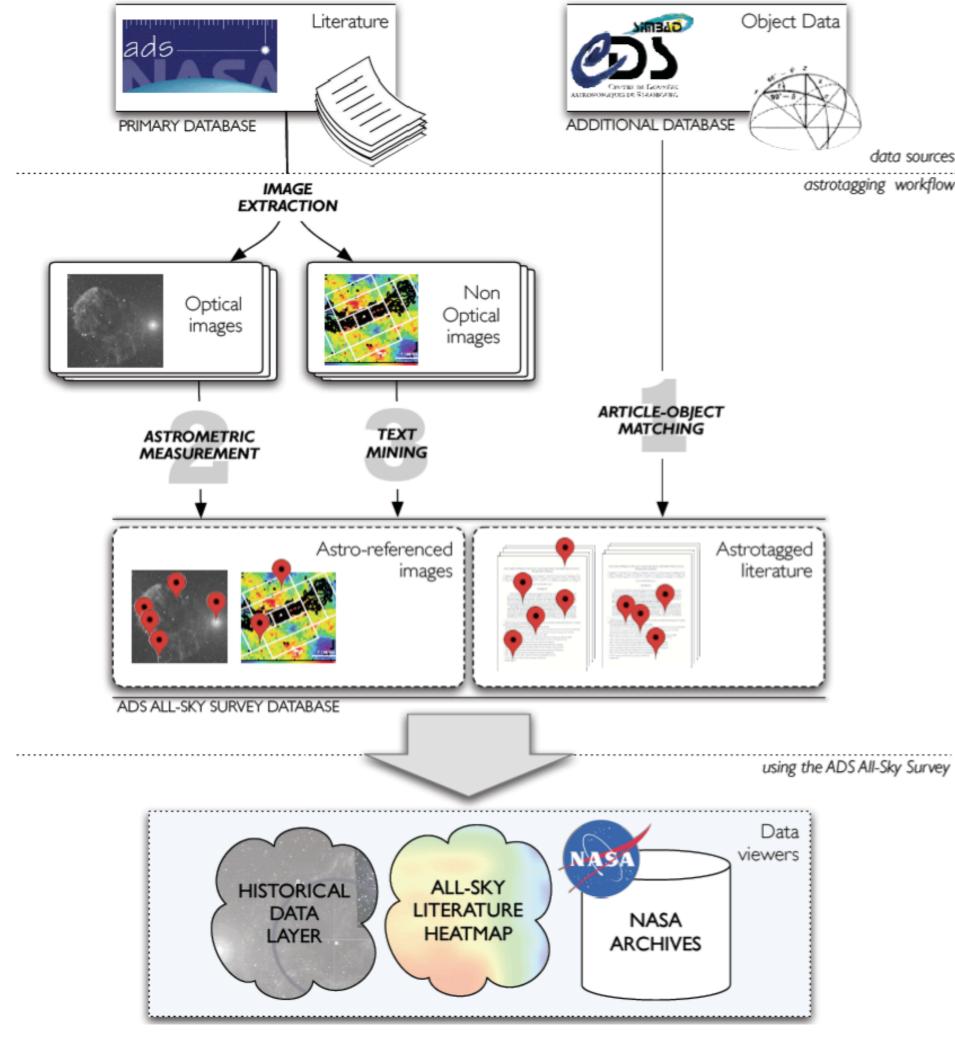


Yes, this is Google Sky...WWT version coming soon!

## c.2012

The ADS All-Sky Survey

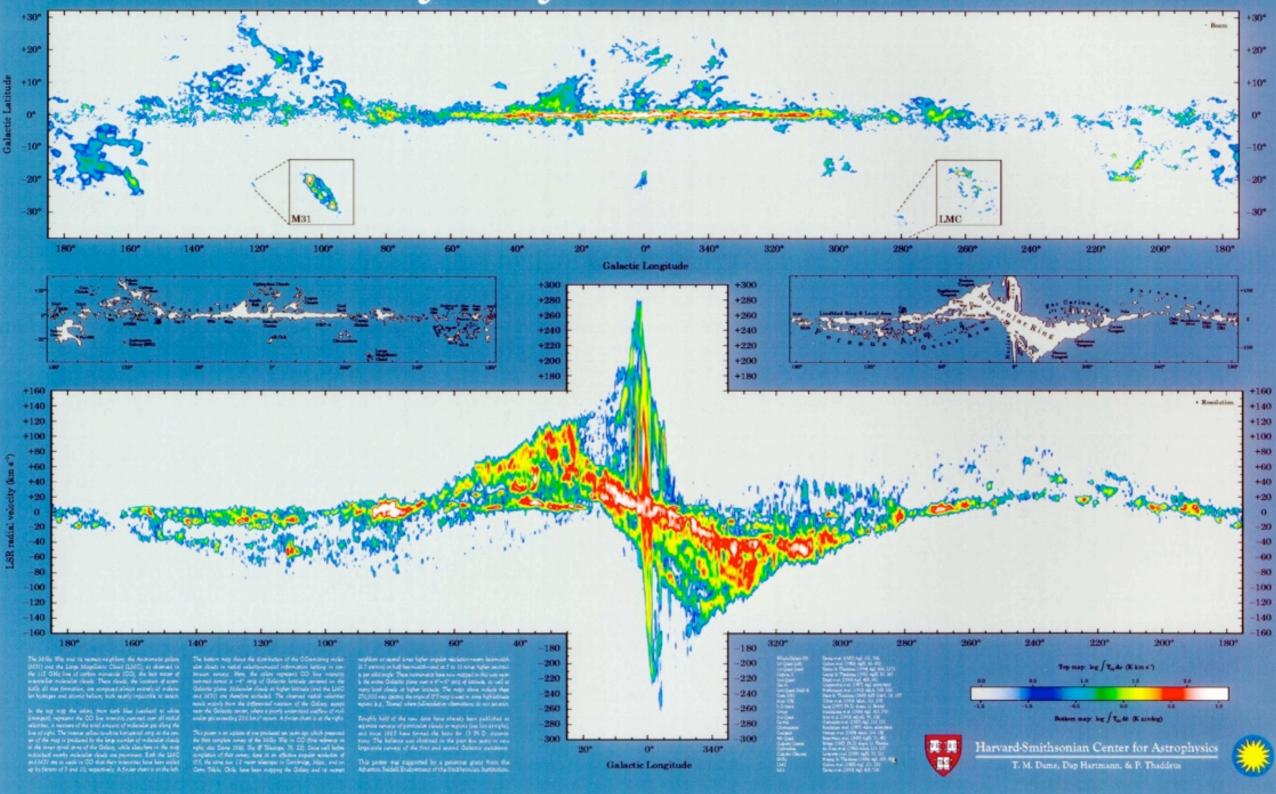
> A. Goodman (**CfA**) A. Muench (CfA) A. Pepe (CfA) with A. Accomazzi (CfA), A. Conti (**STScI**), R. Davé (CfA) T. Boch (**CDS**), J. Fay (**MSR**), D. Hogg (**NYU**)



# The Future universe3d.org



## The Milky Way in Molecular Clouds



Dame, Hartman & Thaddeus 1997



3D Viz made with VolView

## AstronomicalMedicine@







ABOUT PEOPLE

PROJECTS

PUBLICATIONS PRESENTATIONS

#### Projects



#### Seamless integration of scientific data and literature

Astronomical data artifacts and publications exist in disjointed repositories. The conceptual relationship that links data and publications is rarely made explicit. In collaboration with **ADS** and **ADSIabs**, and through our work in conjunction with the Institute for Quantitative Social Science (IQSS), we are working on developing a platform that allows data and literature to be seamlessly integrated, interlinked, mutually discoverable.



#### Astronomy Dataverse

Astronomers use, peruse and produce vast amounts of scientific data. Making these data publicly available is important because it supports the reproducibility of results, and ensures their long term preservation and reuse. While raw astronomical data are normally stored and made public available via large-scale archives, reduced data are often left out entirely from both astronomical archives and related publications.

In a pilot study in 2011, we are evaluating the Dataverse, an open data archive hosted by Harvard University and managed by the Institute for Quantitative Social Science (IQSS), as a project-based repository for the storage, access, and citation of reduced astronomical data. We have interviewed a set of 10 astronomers about their needs, and the prototype CfA Dataverse is now online.

## World astron

WorldWide Telescope (WWT)

WorldWide Telescope provides a rich contextual visualization environment for astronomical data. Our group collaborates with the WWT Team at Microsoft Research both to enrich WWT for use in research as well as in teaching. On the research end, we seek to integrate WWT "Seamlessly" with VAO-sponsored projects, as well as with ADS Labs. On the teaching end, we founded and now run the WorldWide Telescope Ambassadors outreach effort. Events Twitter Links

#### Latest news

augustmuench: Farmers' Almanac foresees a rough winter ahead & dowser finds huge body of dirty water under the Longfellow bridge http://t.co/DwL2mMT

#### albertoconti: RT

@james\_s\_bullock: Why doesn't the History Channel just change its name to the Bigfoot Lover's Pseudoscience Channel?http://j.mp/rlKp4C

albertoconti: RT @johnmaeda: "Not everything knowable can be articulated in propositional form." http://t.co/ZYD43ER

albertoconti: RT @sarahkendrew: looks awesome! >> @astrobetter: New Post: iObserve: The Astronomical Observing App We've Been Waiting For http:/ ...

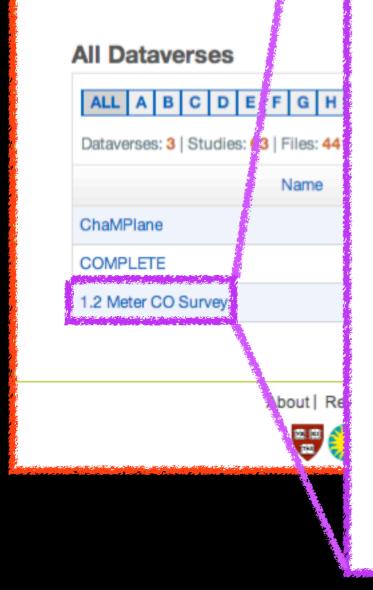
augustmuench: my post `on open science and anonymous peer review`

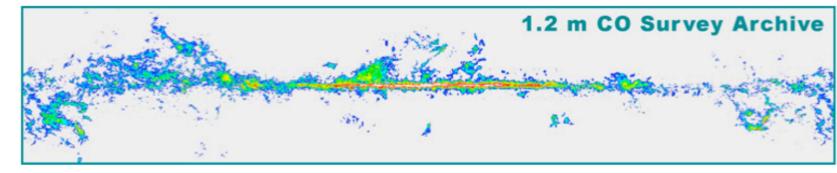


#### Astronomy Dataverse N

Search User Guides Report Issu

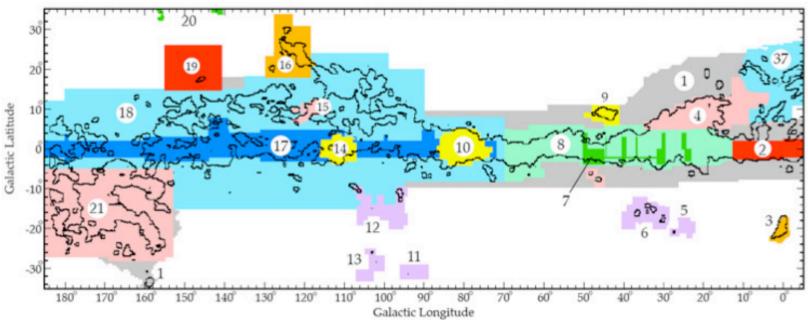
This is a prototype for an Astronomy c Astrophysics, the ADS, the Welbach L

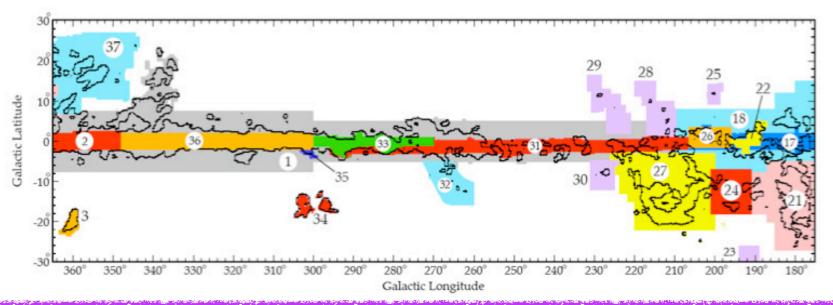




All Astronomy Dataverses > 1.2 Meter CO Survey Dataverse	POWERED BY THE <b>Dataverse</b> PROJECT
Search User Guides Report Issue	Log In Create Account

The whole-Galaxy CO survey presented in Dame et al. (2001) is a composite of 37 separate surveys that are described and numbered in Table 1 of the paper. The data from most of these surveys can be accessed by clicking on the survey number in the map below, which is Figure 1 from the paper. Larger composites of these individual surveys are available from the link below. ... more >>







#### aagie Alyssa Goodman

If you know of "viewers" for maps of the Galaxy or Universe online, add them here: tinyurl.com/universe3d, and pass on this link too. 27 Sep



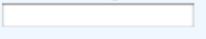
*Pinky*: "Gee, Brain, what do you want to do tonight?" *The Brain*: "The same thing we do every night, Pinky—try to take over the world!"

#### **Online Maps of the Galaxy & the Universe**

This form is being used to collect information about online resources that offer data-driven views of the Milky Way or the Universe beyond. Ultimately, the information here will be used to populate a new "aggregator" service at <u>universe3d.org</u>.

\* Required

#### Name of Site/Page/Service \*



URL: \* enter the main URL for the site

#### Alternate URL:

if there is more than 1 URL associated with the service, enter it, or a note about it, here

Contact email or URL for site creator/maintainer, if known:

Your email address if you're willing to elaborate on your answers here, should we have questions..

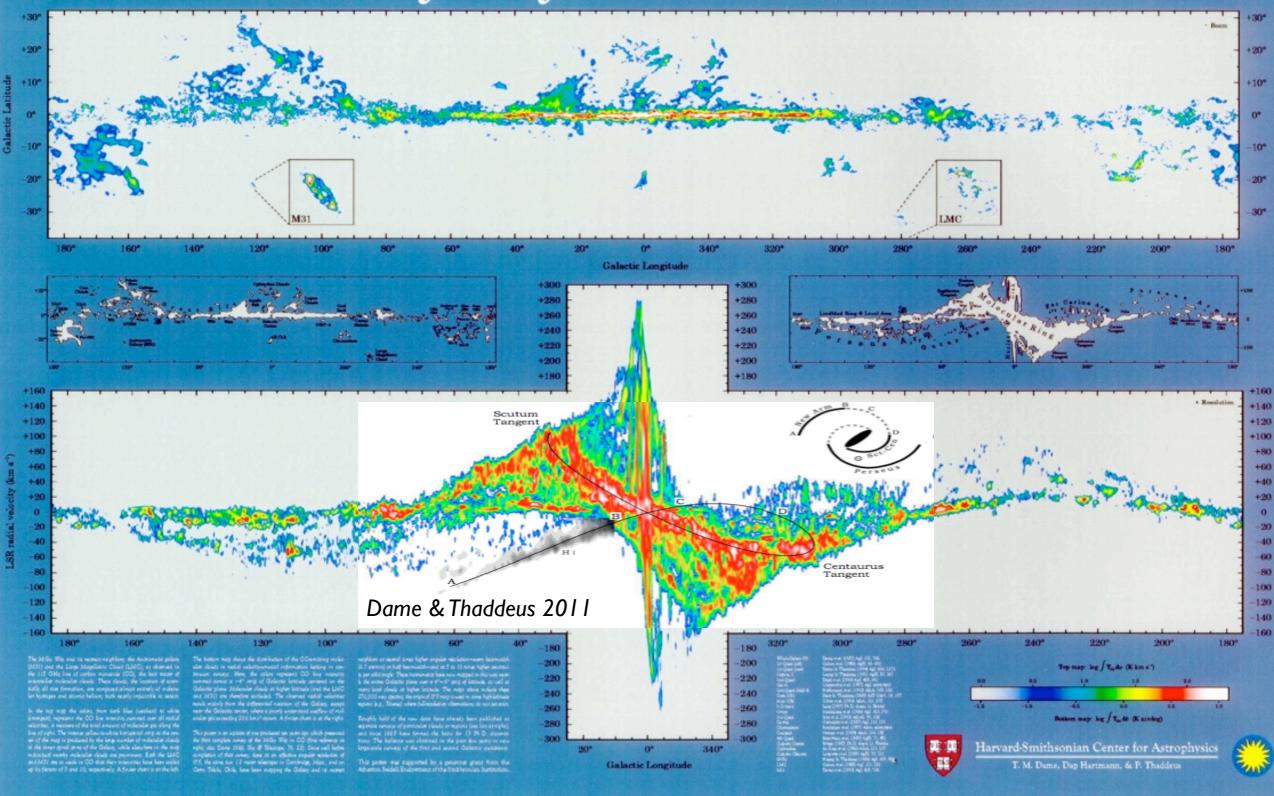
(Optional, but helpful to us!)

If you'd like to, please describe what you think this site is useful for as it stands...

#### Software tools used

Please describe, in free text format, anything you know about how the site works, from a software point of view.

## The Milky Way in Molecular Clouds



Dame, Hartman & Thaddeus 1997

# Tools for Taking over the World:

WWT + more SAMP-enabled tools + Linked Views (+...)

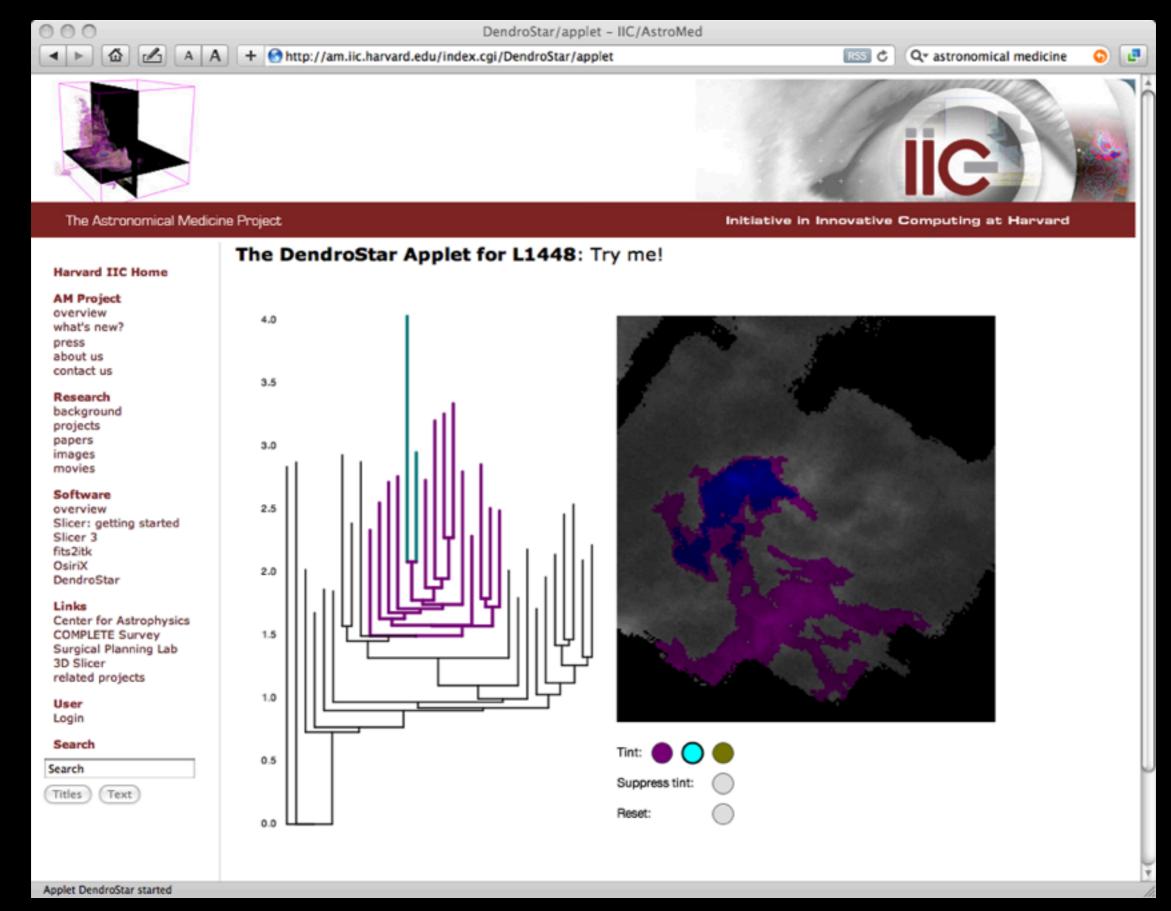
## Contextual, High-Dimensional View



Flat, Text-Based View

LX:

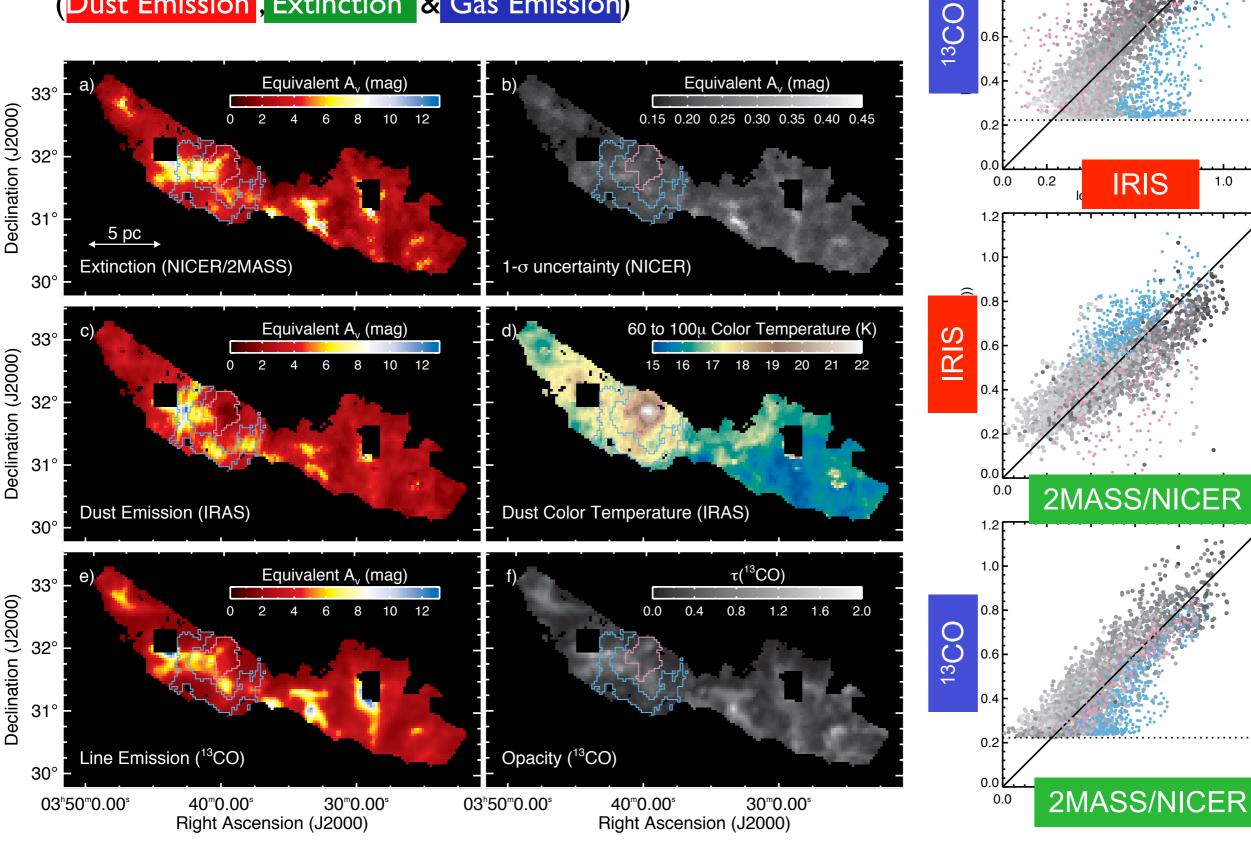
Jan Vermeer. The Astronomer. (1668)



<u>http://am.iic.harvard.edu/index.cgi/DendroStar/applet</u> Dendrogram Algorithm by Erik Rosolwosky;Applet by Douglas Alan

## **COMPLETE** Perseus Column Density

(Dust Emission, Extinction & Gas Emission)



figures: Goodman, Pineda & Schnee 2009 cf. Schnee et al. 2005, 2006, 2008; Pineda et al. 2008

1.2

0.0

0.0

0.0

1.2

1.2

IRIS

1.0

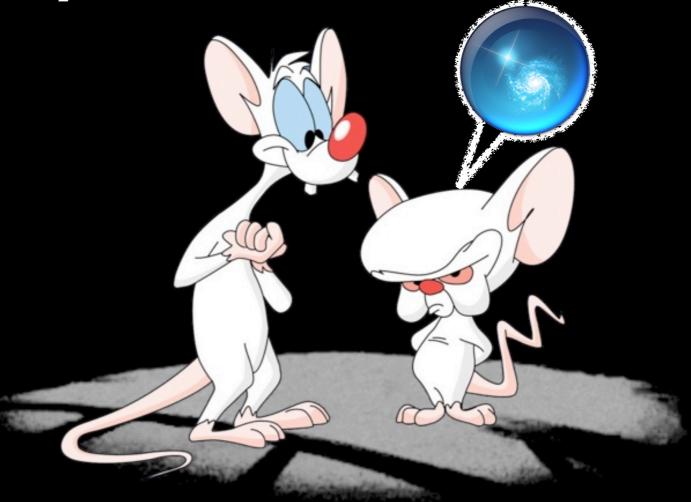
1.2

1.2

1.0

0.8

# Seamless Astronomy Enabled by WWT



## Alyssa A. Goodman Harvard-Smithsonian Center for Astrophysics

From: Abstract Service <ads@cfa.harvard.edu> Subject: myADS Notification (Astronomy database) Date: March 23, 2010 12:19:23 AM EDT To: Alyssa Goodman

ads	myADS Personal Notification Service for Alyssa Goodman Tue Mar 23 00:19:23 2010 Astronomy database	
ADS Main Queries Astronomy GOODMAN, ALYSSA - Citations: 3310 (total 4002)	Favorite Authors - Recent Papers No new articles found	
Physics arXiv e-prints FAQ 2010NewA15444K: Karatas,- intrinsic-colour calibration for uv photometry 2010MNRAS.403.1054D: Dabri	+: New /by-beta inghausen,+: <u>2010A&amp;A511A90B</u> : Breddels,+: Distance	
What's new         Mass loss and expansion of ultradwarf galaxies through gas expansion           Current Tables of Contents         Stellar evolution for top-heavy smass functions           Astronomical Journal Astronomy &         Clouds: Implementation and Complementation	From: Kayak Alert <alert@kayak.com> Subject: Your KAYAK Fare Alert: Boston (BOS) &gt; Munich (MUC) Date: March 26, 2010 3:52:30 AM EDT To: Alyssa Goodman Reply-To: Kayak Alert <alert@kayak.com></alert@kayak.com></alert@kayak.com>	
Astrophysics Astrophysics Astrophysics Supplements Astrophysics Supplements Astrophysics Supplements Astrophysics Supplements Astrophysics Supplements Astrophysics Supplements Astrophysics Supplements Astrophysics Supplements Astrophysics Supplements Astrophysics Supplements Astrophysics Supplements Astrophysics Supplements Astrophysics Supplements Astrophysics Supplements Astrophysics Supplements Astrophysics Supplements Astrophysics Supplements Astrophysics Astrophysics Supplements Astrophysics Supplements Astrophysics Astroph	K A Y A K	Fare Alert
Astrophysical Journal 2010ApJ712.1137K: Kauffma	Flight Deals Hotel Deals Vacation & Package Deals Cruis	e Deals Top Deals
The "travel" analogy	Boston (BOS) to Munich (MUC), round trip Saturday, Jun 12 to Saturday, Jun 19	Don't miss out on big savings
seems to resonate	\$1649       \$0 since yesterday         per person, economy, nonstop       \$1500	Get deals by email
best	Check Now 10 Days Ago Today	Low Fares* found from Boston (BOS) to: \$99+ <u>Baltimore</u>
	Today's best fares - cheapest 3 of 3 airlines (Mar 26, 3:51a EDT)	\$144+ <u>Orlando</u> \$146+ <u>Washington</u>

All Results

<u>\$1649</u>

nonstop

1 eton

Lufthansa

<u>\$1649</u>

Multiple

Airlines

<u>\$1656</u>

\$152+ Atlanta

\$184+

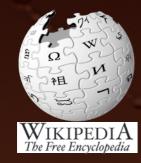
<u>Fort</u> Lauderdale

United

<u>\$1649</u>

## Literature







### Blogs, Wikis, etc.

## "Seamless Astronomy" (Tools)

CLUMPFIND segmentation

LETTERS

0

Figure 2 | Comparison of the 'dendrogram' and 'CLUMPFIND' feature

identification algorithms as applied to <sup>13</sup>CO emission from the L1448 region of Perseus. a. 3D visualization of the surfaces indicated by colours in

the dendrogram shown in c. Purple illustrates the smallest scale self-

gravitating structures in the region corresponding to the leaves of the

dendrogram: pink shows the smallest surfaces that contain distinct self-

data cube containing all the significant emission. Dendrogram branches corresponding to self-gravitating objects have been highlighted in yellow

the virial parameter is less than 2. The x-y locations of the four 'selfgravitating' leaves labelled with billiard balls are the same as those shown in Fig. 1. The 3D visualizations show position–position–velocity (p-p-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

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

Four years before the advent of CLUMPFIND, 'structure trees'

were proposed as a way to characterize clouds' hierarchical structure

front (-0.5 km s<sup>-1</sup>) to back (8 km s<sup>-1</sup>)

(Supplementary Fig. 1).

gravitating leaves within them; and green corresponds to the surface in the

over the range of  $T_{\rm mb}$  (main-beam temperature) test-level values for which

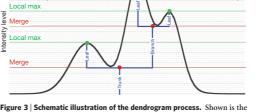
## Data











NATURE Vol 457 1 January 2009

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  $(p-p-\nu)$  data cube into an easily visualized representation called a 'dendrogram'<sup>10</sup>. Although well developed in other data-intensive fields<sup>11,12</sup>, 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

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

A dendrogram of a spectral-line data cube allows for the estimation

of key physical properties associated with volumes bounded by iso-

surfaces, such as radius (*R*), velocity dispersion ( $\sigma_v$ ) and luminosity (*L*). The volumes can have any shape, and in other work<sup>14</sup> 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  $M_{\text{lum}} = X_{13CO}L_{13CO}$ , where  $X_{13CO} = 8.0 \times 10^{20} \text{ cm}^2 \text{ K}^{-1} \text{ km}^{-1} \text{ s}$ (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,  $\alpha_{obs} = 5\sigma_v^2 R/GM_{lum}$ 

In principle, extended portions of the tree (Fig. 2, yellow highlighting)

where  $\alpha_{obs} < 2$  (where gravitational energy is comparable to or larger

than kinetic energy) correspond to regions of p-p-v space where self-

gravity is significant. As  $\alpha_{obs}$  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 fields16, its measured

value should only be used as a guide to the longevity (boundedness) of

'merger trees' are being used with increasing frequency<sup>13</sup>. 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

online) and a sorted dendrogram (Fig. 2c).

any particular feature.

Local max

Test level

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 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.



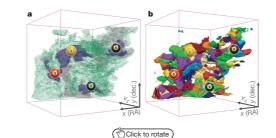
**Disclaimer**:This slide shows key excerpts from within the astronomy community & excludes more general s/w that is used, such as Papers, Zotero, Mendeley, EndNote, graphing & statistics packages, data handling software, search engines, etc.

©2009 Macmillan Publishers Limited. All rights reserved

Data in *Literature* 

#### LETTERS

#### NATURE Vol 457 |1 January 2009



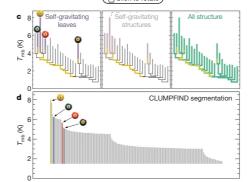


Figure 2 | Comparison of the 'dendrogram' and 'CLUMPFIND' featureidentification algorithms as applied to <sup>13</sup>CO emission from the L1448 region of Perseus, a. 3D visualization of the surfaces indicated by colours in the dendrogram shown in c. Purple illustrates the smallest scale selfgravitating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

the printed version, the front face of each 3D cube (the 'home' view in the interactive online version) corresponds exactly to the patch of sky shown in Fig. 1, and velocity with respect to the Local Standard of Rest increases from front ( $-0.5 \text{ km s}^{-1}$ ) to back (8 km s<sup>-1</sup>). 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

be overly dependent on input parameters. By tuning CLUMPFIND's two free parameters, the same molecular-line data set<sup>6</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

64

©2009 Macmillan Publishers Limited. All rights reserved

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'<sup>10</sup>. Although well developed in other data-intensive fields<sup>11,12</sup>, 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 frequency<sup>13</sup>.

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 ( $\sigma_v$ ) and luminosity (1.) The volumes can have any shape, and in other work<sup>14</sup> 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  $\hat{X}_{13\text{CO}} = 8.0 \times 10^{20} \text{ cm}^2 \text{ K}^{-1} \text{ km}^{-1} \text{ s}$ (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,  $\alpha_{obs} = 5\sigma_v^2 R/GM_{lum}$ In principle, extended portions of the tree (Fig. 2, vellow highlighting) where  $\alpha_{obs} < 2$  (where gravitational energy is comparable to or larger than kinetic energy) correspond to regions of p-p-v space where selfgravity is significant. As  $\alpha_{obs}$  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 fields16, its measured value should only be used as a guide to the longevity (boundedness) of any particular feature.

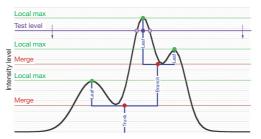


Figure 3 | Schematic illustration of the dendrogram process. Shown is the construction of a dendrogram from a hypothetical one-dimensional emission profile (black). The dendrogram (blue) can be constructed by 'dropping' a test constant emission level (purple) from above in tiny steps (exaggerated in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