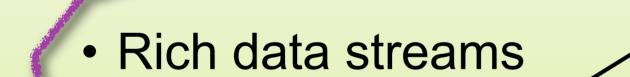
## Cyberinfrastructure meets Cyberlearning

 Infrastructure to support data- and informationintensive distributed work



Scientific learning

Scientific work

## "Seamless Astronomy" & the WorldWide Telescope Alyssa A. Goodman

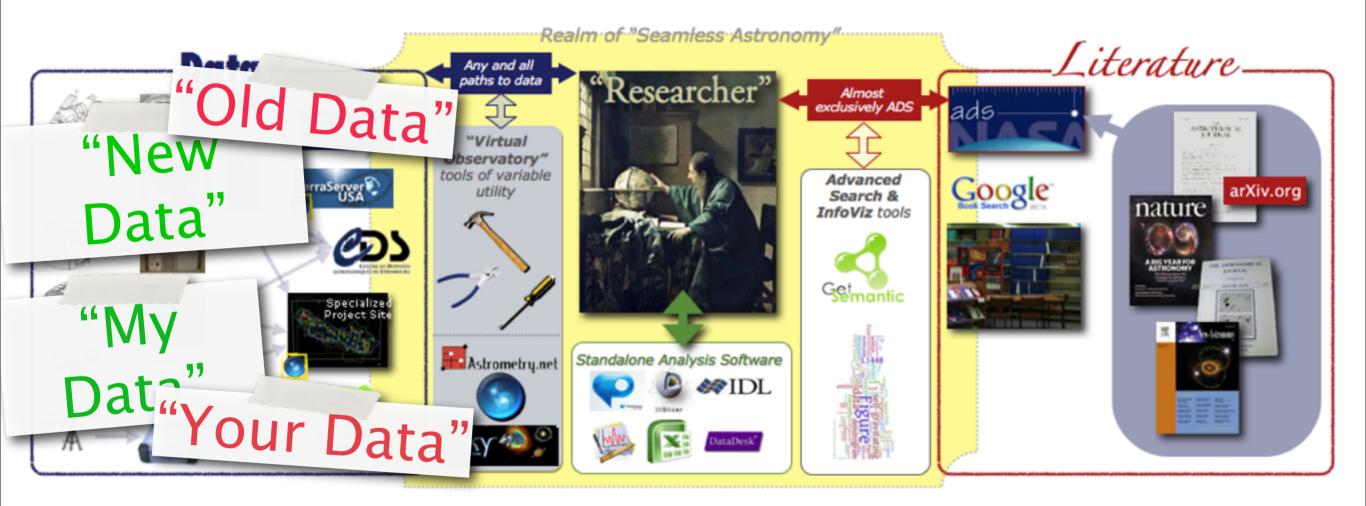
Harvard-Smithsonian Center for Astrophysics Initiative in Innovative Computing @ Harvard

### Collaborators

Harvard-Smithsonian Center for Astrophysics & SEAS: Alberto Accomazzi, Eli Bressert, Douglas Burke, Rahul Davé, Pepi Fabbiano, Michael Kurtz, Gus Muench, Pavlos Protopapas Massachusetts General Hospital: Tim Clark & Sudeshna Das Microsoft Research: Jonathan Fay, Curtis Wong RPI: Jim Hendler & Deborah McGuinness STScl:Alberto Conti & Carol Christian UCLA: Christine Borgman

### Seamless Astronomy

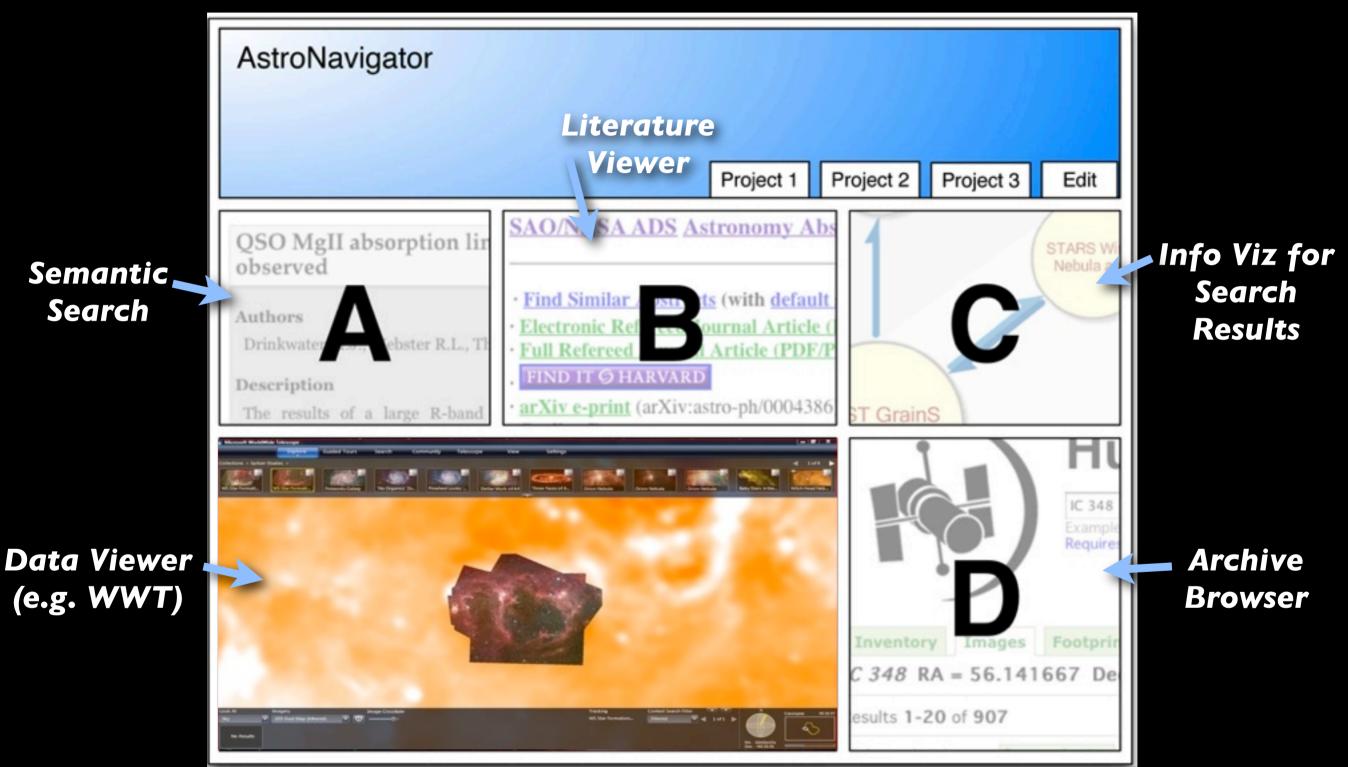
www.cfa.harvard.edu/~agoodman and worldwidetelescope.org



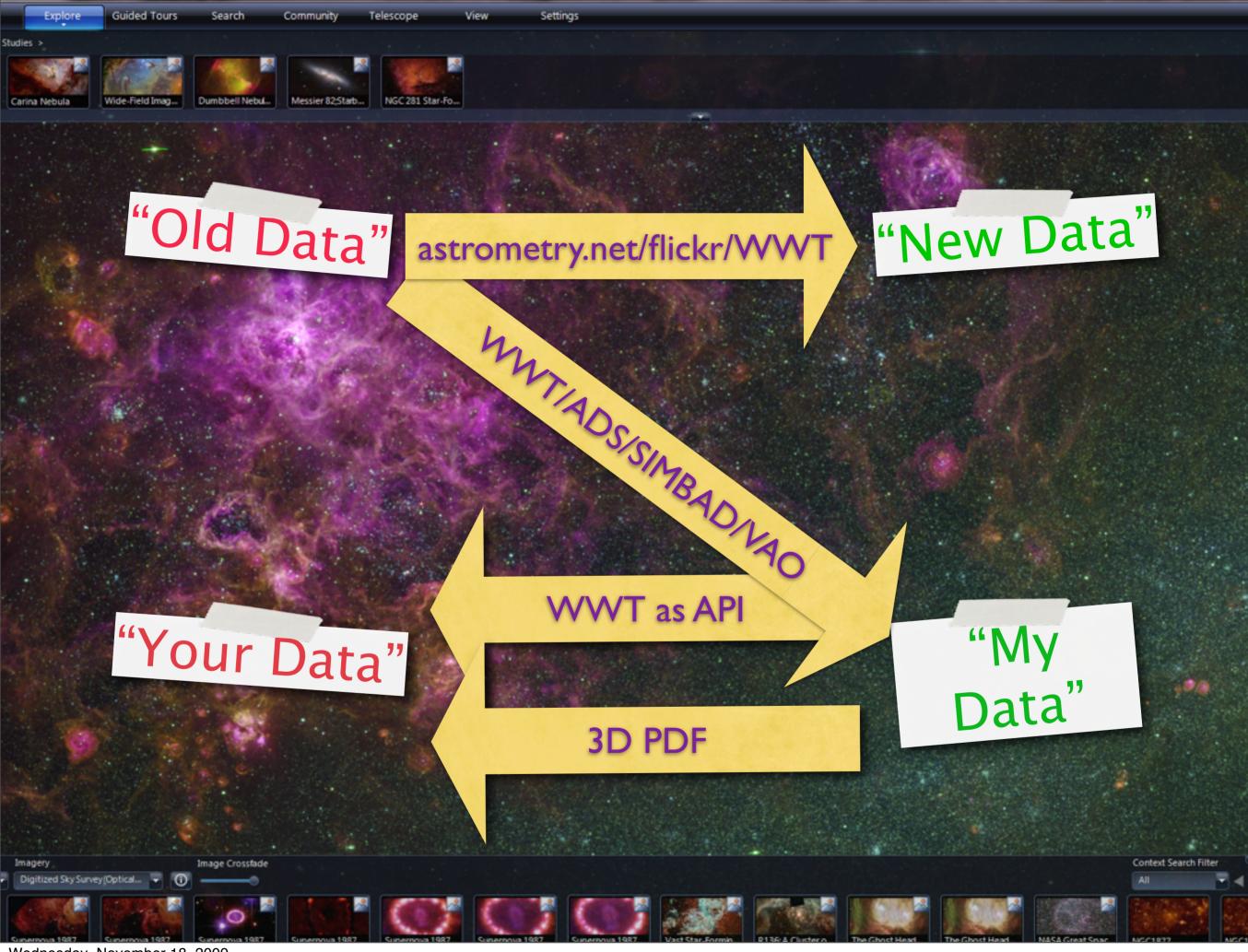
## Astronomers & the V.O. c. 2006

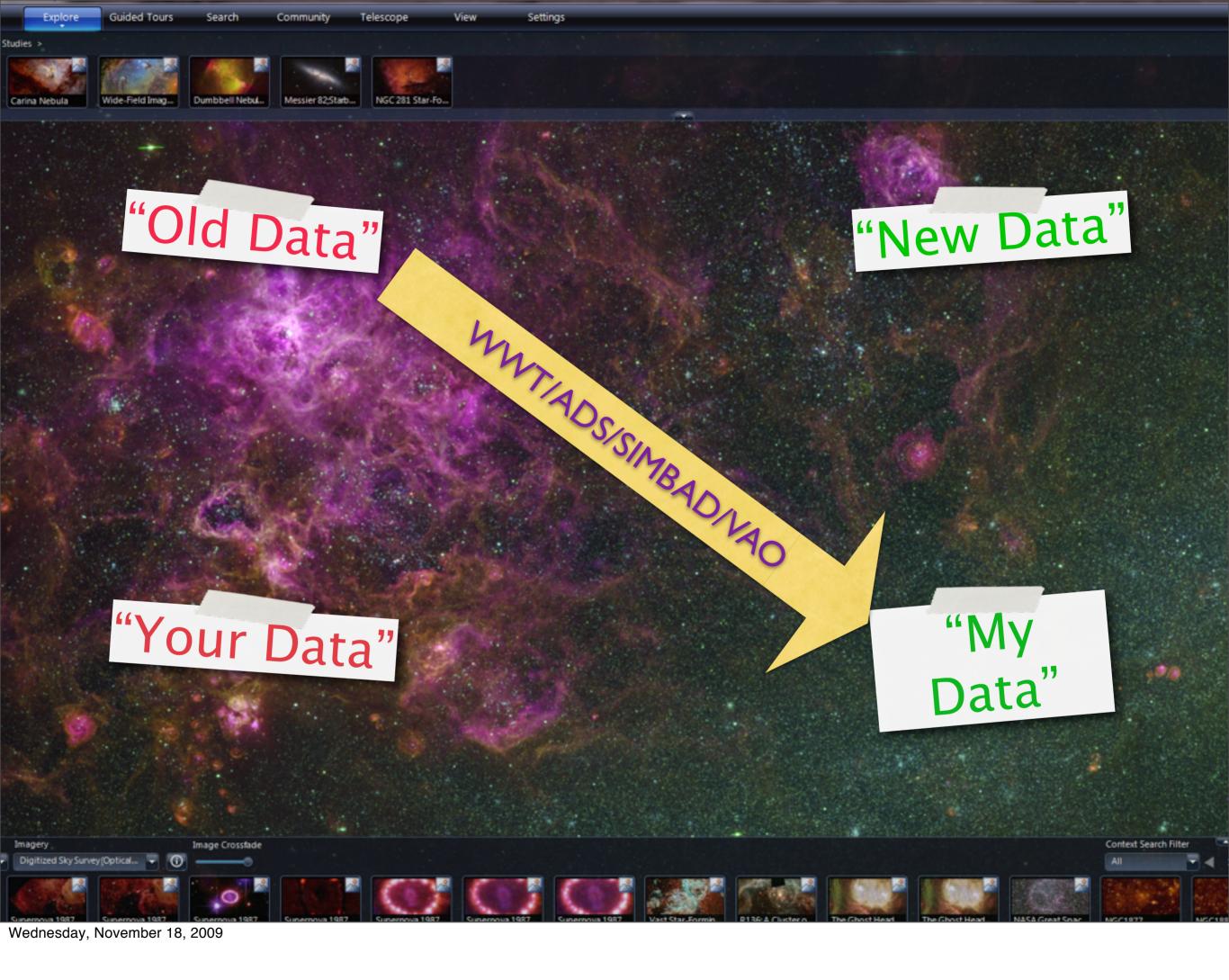


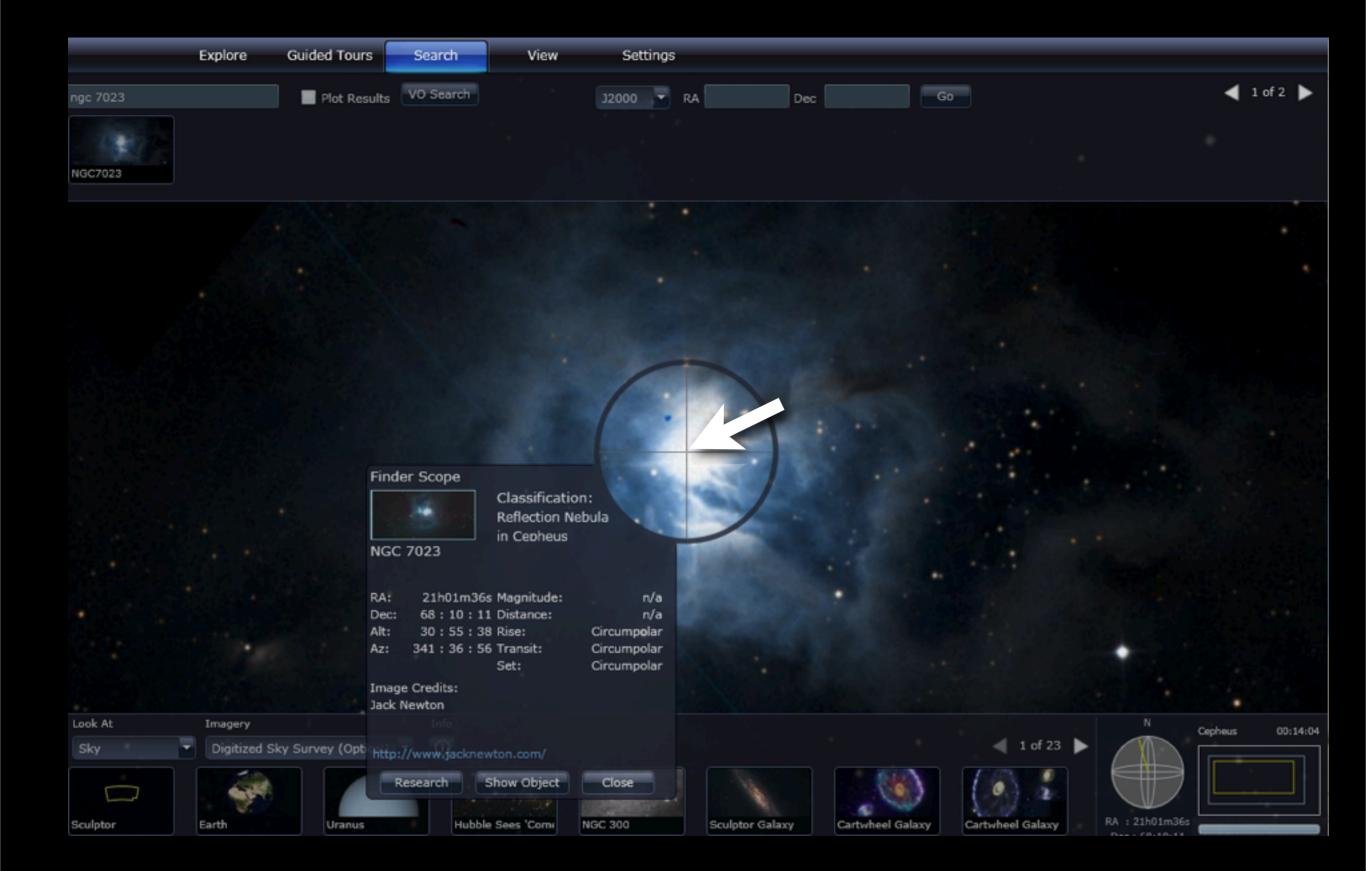
# "Seamless Astronomy"

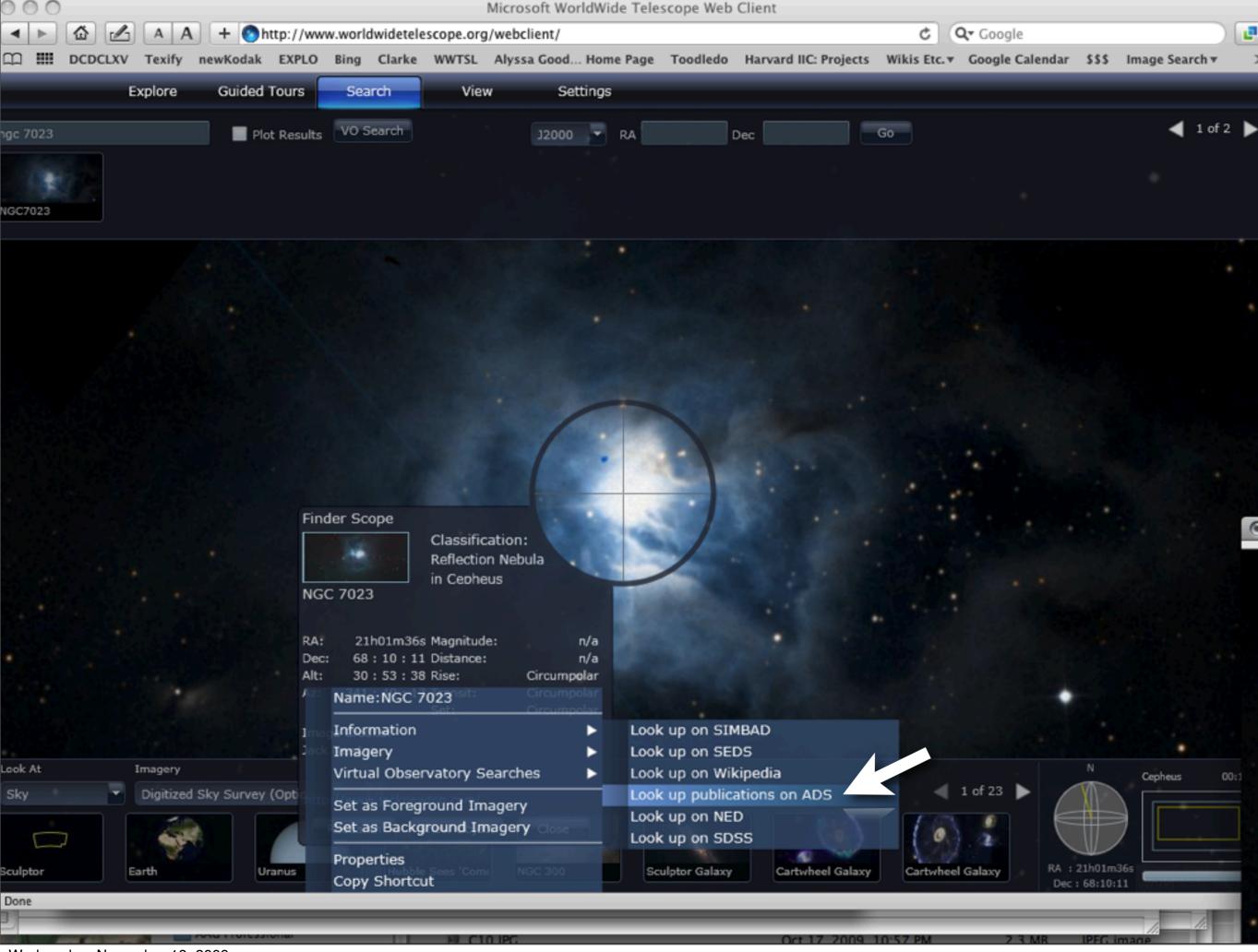


Mockup based on work of Eli Bressert, excerpted from NASA AISRP proposal by Goodman, Muench, Christian, Conti, Kurtz, Burke, Accomazzi, McGuinness, Hendler & Wong, 2008







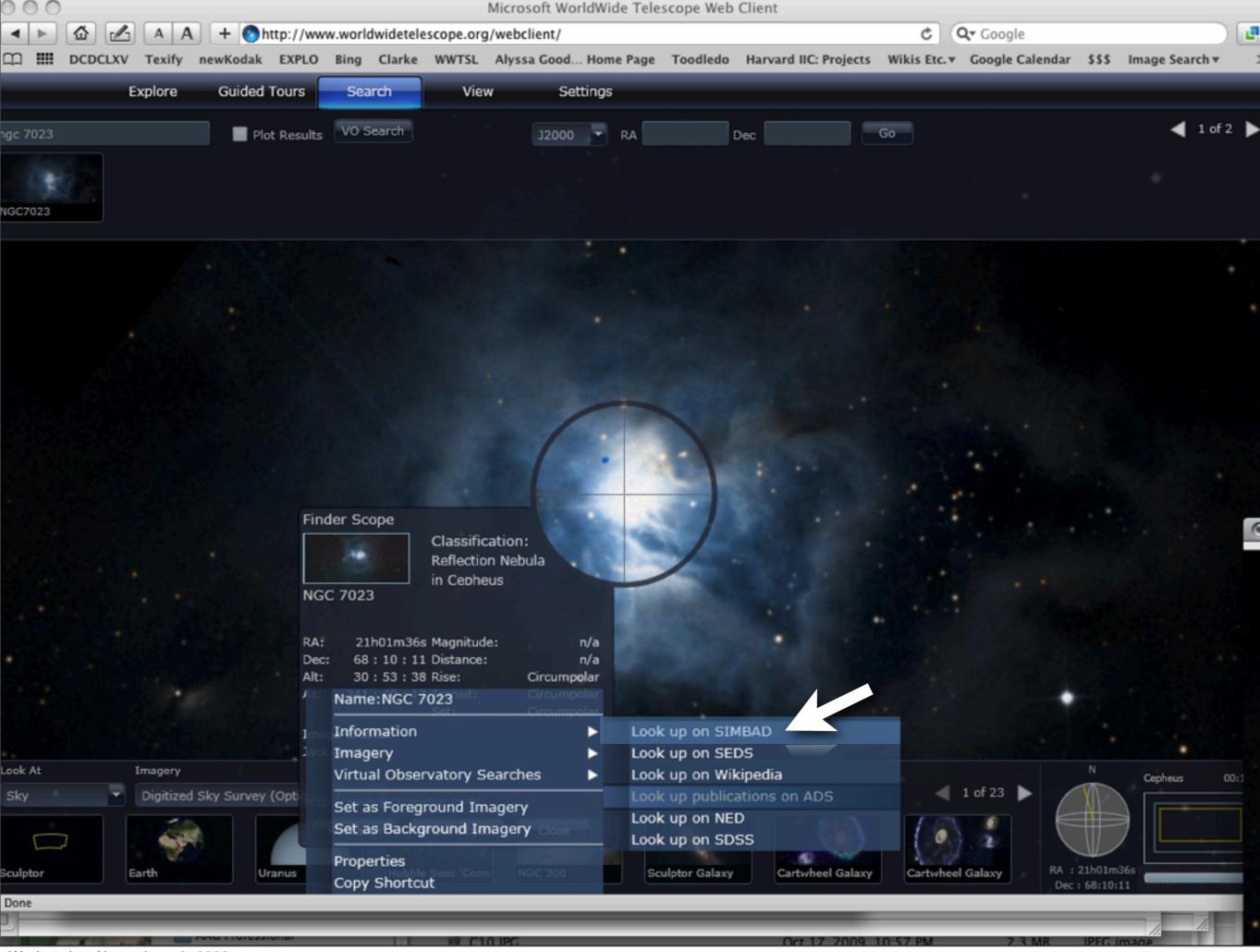


00	0.0								Obje	ect Quer	ry Resu	lts											
•	▶ 🙆 🖌	AA	+ Ade h	ttp://adsa	bs.harva	rd.edu	J/cgi-bin/	abs_co	nnect?db	_key=AS	T∼_	query=1	'ES&obje	ct=NGC%	207 🖪	55 C	Q. (	Google					
m	DCDCLXV	Texify I	newKodak	EXPLO	Bing C	larke	WWTSL	Alyss	a Good I	Home Pa	ge To	odledo	Harvard	l IIC: Proje	ects W	Vikis Etc	.▼ Go	ogle Ca	lendar	\$\$\$	Image Se	arch 🔻	X
SA	O/NASA Ast	rophys	ics Data	System	n (AD)	S)										a	igoodm	an@cfa	a.harva	rd.edu	my Accou	int   <u>Sign</u>	off
	ery Results f																			Go	to botto	m of pag	<u>e</u>
Reti	rieved 200 abstra	acts, starti	ing with n	umber 1.	Total n	umber	selected:	393.											So	rt option	s	4	
#	Bibcode Authors				Score Title		Date		List of L Access C		Help												
1	⊟ 2009ApJ7 Myers, Philip C		М		1.000 Filam		08/2009 y Structur		AZE tar-formi				<u>R</u> <u>C</u>	<u>S</u>	U								
2	2009ApJ7 Desai, Vandana LeFloc'h, Emeri Brown, Michae Jannuzi, Buell T coauthors	i; Soifer, l ic; Armus I J. I.; Br	B. T.; Dey s, Lee; Bra odwin, Ma	and, Kate ark;			08/2009 vcyclic A		AZE ic Hydrod			on from	<mark>RC</mark> z≈2U	<mark>S</mark> LIRGs	U								
3	2009MNRA Nutter, D.; Stan Thompson, D.				1.000 The i		07/2009 conditions		AZE			X. Akar	R i mappi	<mark>S</mark> ng of an	U extern	ally hea	ated pr	e-stella	ar core	;			-
4	■ 2009A&A Boersma, C.; Pe Hernández, N. I Verhoeff, A. P.; Waters, L. B. F.	eters, E.; L.; van de ; Tielens,	; Martín- er Wolk, C , A. G. G.		1.000 A spa		07/2009 udy of the		AZE IR emiss		ures in	four He	R rbig Ae	<u>S</u> /Be stars	U s								-
5	2009MNRA Hernán-Caballe Hatziminaoglou Rowan-Robinso Farrah, D.; Lon	ro, A.; Po , E.; Afo on, M.; R	érez-Fourr nso-Luis, igopoulou	A.; 1, D.;	1.000 Mid-i		05/2009 od spectro		AZE of infrar			alaxies a	$\frac{\mathbf{R}}{\mathbf{C}} = \mathbf{C}$		U								-

000	PV Cephei: Young Star Caught Speeding?											
< > & Z	A A + Mattp://adsabs.harvard.edu/abs/2004ApJ608831G	Ċ	Q <del>.</del> Google									
	xify newKodak EXPLO Bing Clarke WWTSL Alyssa Good Home Page Toodledo Harvard IIC: Projects	Wikis Etc. •	Google Calendar	\$\$\$ Image Sear		>>						
		390	odman@cfa.harvar	dedul my Account	Sien o	ar Å						
		age	ounanecialiarvar	uzuu i <u>my Account</u>	Signo	" N						
SAO/NASA ADS Astronomy Abstract Service												
						- 11						
Find Challen Abstra	ete (mith defeatte estime helen)											
Custom Format	cts (with <u>default settings below</u> )					- 11						
	Journal Article (HTML)					- 11						
Electronic Refereed Journal Article (HTML)     Full Refereed Journal Article (PDF/Postscript)												
FIND IT G HARVAR	D					- 11						
<ul> <li>arXiv e-print (arXiv:</li> </ul>	astro-ph/0401486)					- 11						
• <u>On-line Data</u>												
<u>References in the article</u> <u>Citation History</u> )												
Refereed Citations to the Article												
SIMBAD Objects (37)												
Also-Read Articles (Reads History)												
• HEP/Spires Information												
• Translate This Page												
The	DV Canhair Vauna Star Caught Speeding?					ľ						
Title:	PV Cephei: Young Star Caught Speeding?					- 11						
Authors:	Goodman, Alyssa A.; Arce, Héctor G.	Rofe home	ad a day) A D (Calif	and Institute of		- 11						
Affiliation:	AA(Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138; agoodman@ Technology, 1200 East California Boulevard, Pasadena, CA; harce@astro.caltech.edu)	gera.narva	rd.edu), AB(Calif	ornia institute of								
Publication:	The Astrophysical Journal, Volume 608, Issue 2, pp. 831-845. (ApJ Homepage)					- 11						
	06/2004					- 11						
Origin:	UCP					- 11						
ApJ Keywords:	ISM: Herbig-Haro Objects, ISM: Individual: Alphanumeric: HH 315, ISM: Jets and Outflows, Stars: Form	nation. Star	s: Individual: Con	stellation Name: 1	PV	- 11						
	Cephei, Stars: Kinematics	auton, o au				- 11						
Abstract Copyright:	(c) 2004: The American Astronomical Society					- 11						
DOI:	10.1086/383139											
<b>Bibliographic Code:</b>	2004ApJ608831G											
	Abstract											

Three independent lines of evidence imply that the young star PV Cep is moving at roughly 20 km s<sup>-1</sup> through the interstellar medium. The first and strongest suggestion of motion comes from the geometry of the Herbig-Haro (HH) knots in the ``giant" HH flow associated with PV Cep. Bisectors of lines drawn between pairs of knots at nearly equal distances

•



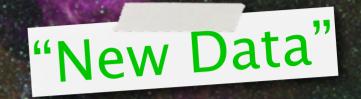
" )	ou	r D	ata	a"				"N					
"	Old	Da	ata	"				Da	ta"				
000						MBAD query				Google			
	A A + 🛋 exify newKodak						-	8F7CD92574727FFC			Search v >>		
(D)			Ø	latalogs	Dictionary	Biblio	Tutorials	Developers			ŕ		
other query Identifier Query guery	Coordinate Criter		bhy Basic guery	SIMBA Script submission	D query res Output options								
Object query : NGC 70	23								C.D.S S	IMBAD4 rel 1.132 - 2009.10.23	CEST21:59:31		
Available data	Basic d	lata	<b>Identifiers</b>	E	lot & images	1	Bibliography	Measure	ments	External archives	Notes		
Other object types:	NGC 7023 Open (galactic) Cluster												
ICRS coord. (ep=2000) : FK5 coord. (ep=2000 eq=2 Fluxes (1) :	000):21 01 36.	+14.1926 (	3 ( - ) [ -	]D- ]D-									
Identifiers (11) :													
NGC 7023 C 2059+679 C1 VDB 139	IRAS	20599+6755 F20599+6755 04.08+14.21			5	A87 CISM 50 C1 235		[BDB2003] G104.06 AAVSO 2044+67	+14.19				
Plots and Images	(plot around)		in previewer)	(Nadin app	<u>let</u>								
	adius 10 ar		A Hersheld	<b>()</b>									
References (371 betwee Simbad bibliographic survey be			stars) and in 198	83 for all other	r objects (outside	the solar system)	ı.			to: 2009			
											3		



Studies >



# "Old Data"



# "Your Data"

mage Crossfad

0

### WWT as API

Settings

## "My Data"

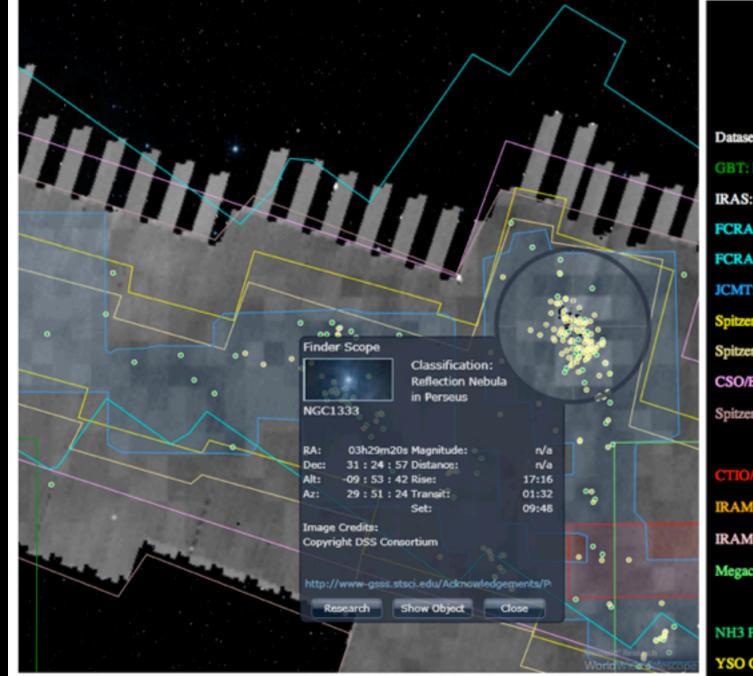
**Context Search Filter** 

Digitized Sky Survey (Optical... 💂

Imagery



0.	COMPLETE Data Coverage Tool																	
A A + Ohttp://www.worldwidetelescope.org/COMPLETE/WWTCoverageTool.htm# C Qr Google																		
1		newKodak	EXPLO	Bing	WWTSL	Alyssa Good Home Page	Toodledo	Harvard IIC: Projects	Wikis Etc. 🔻	Google Calendar	\$\$\$	Image Search *	fbk share	Directories <b>v</b>	ADS Best	RSS (3387) <del>v</del>	BeyondADS <b>*</b>	22



#### **COMPLETE Data Available**

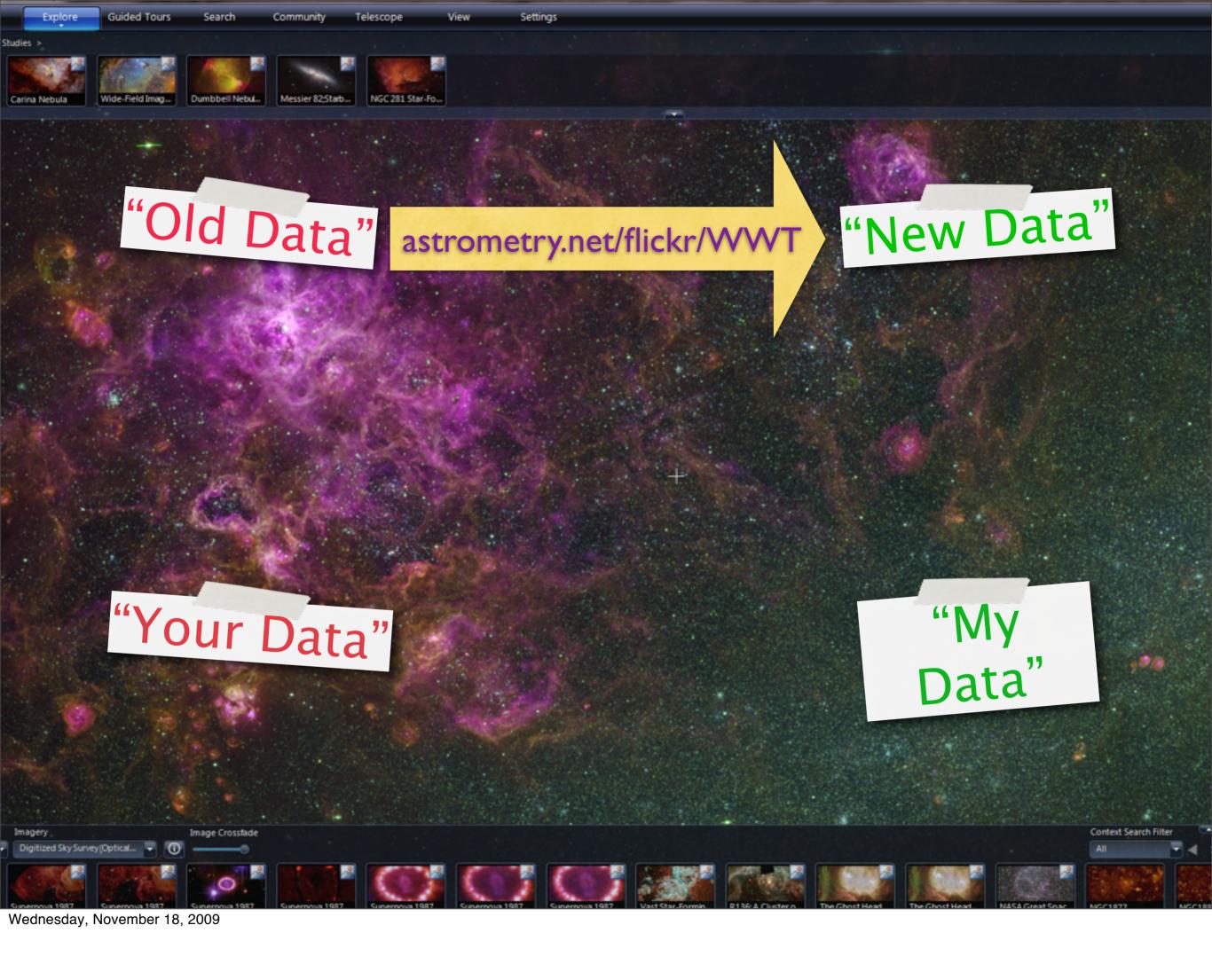
#### Center on Perseus Center on Ophichus Center on Serpens

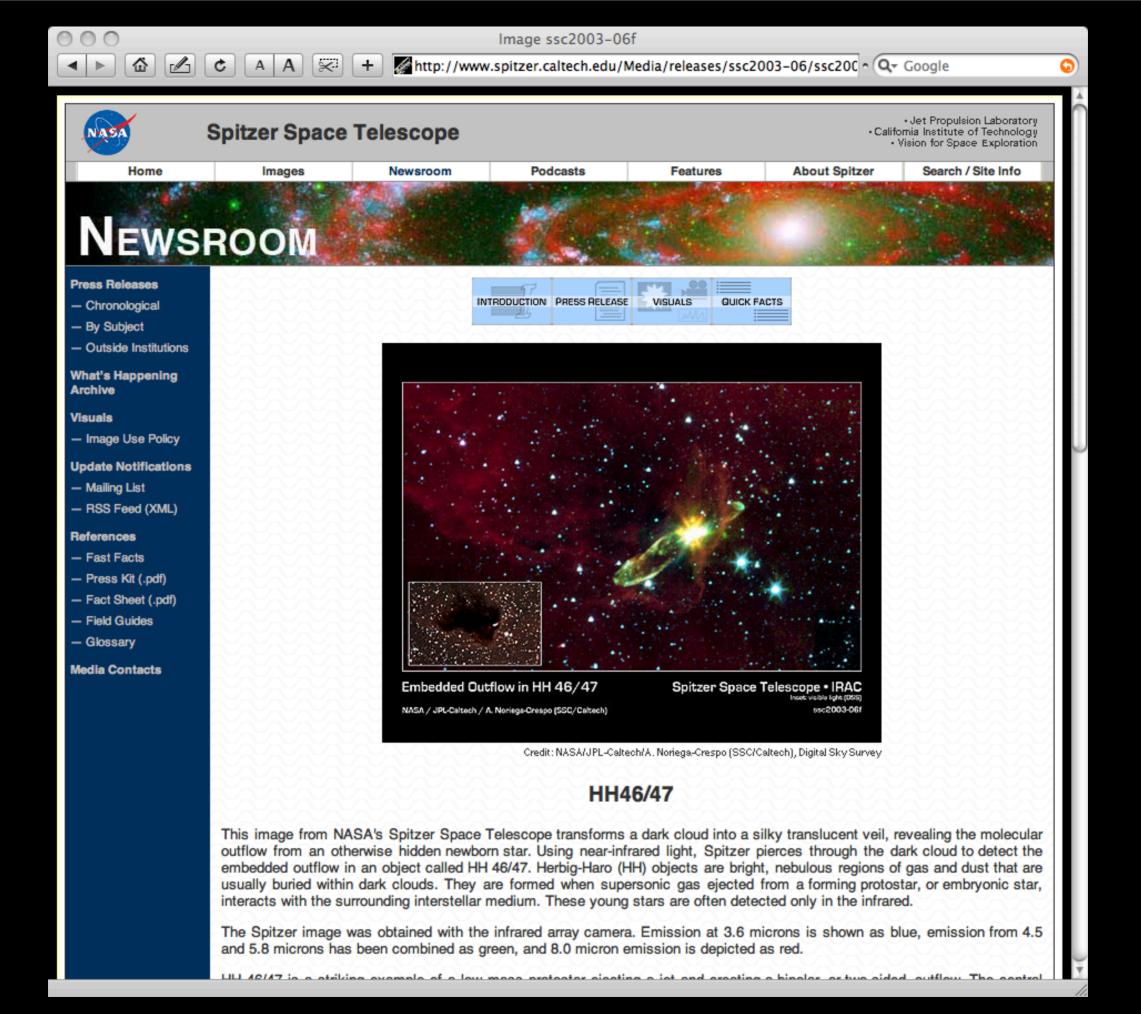
#### Full-Cloud Data (Phase I, All Data Available)

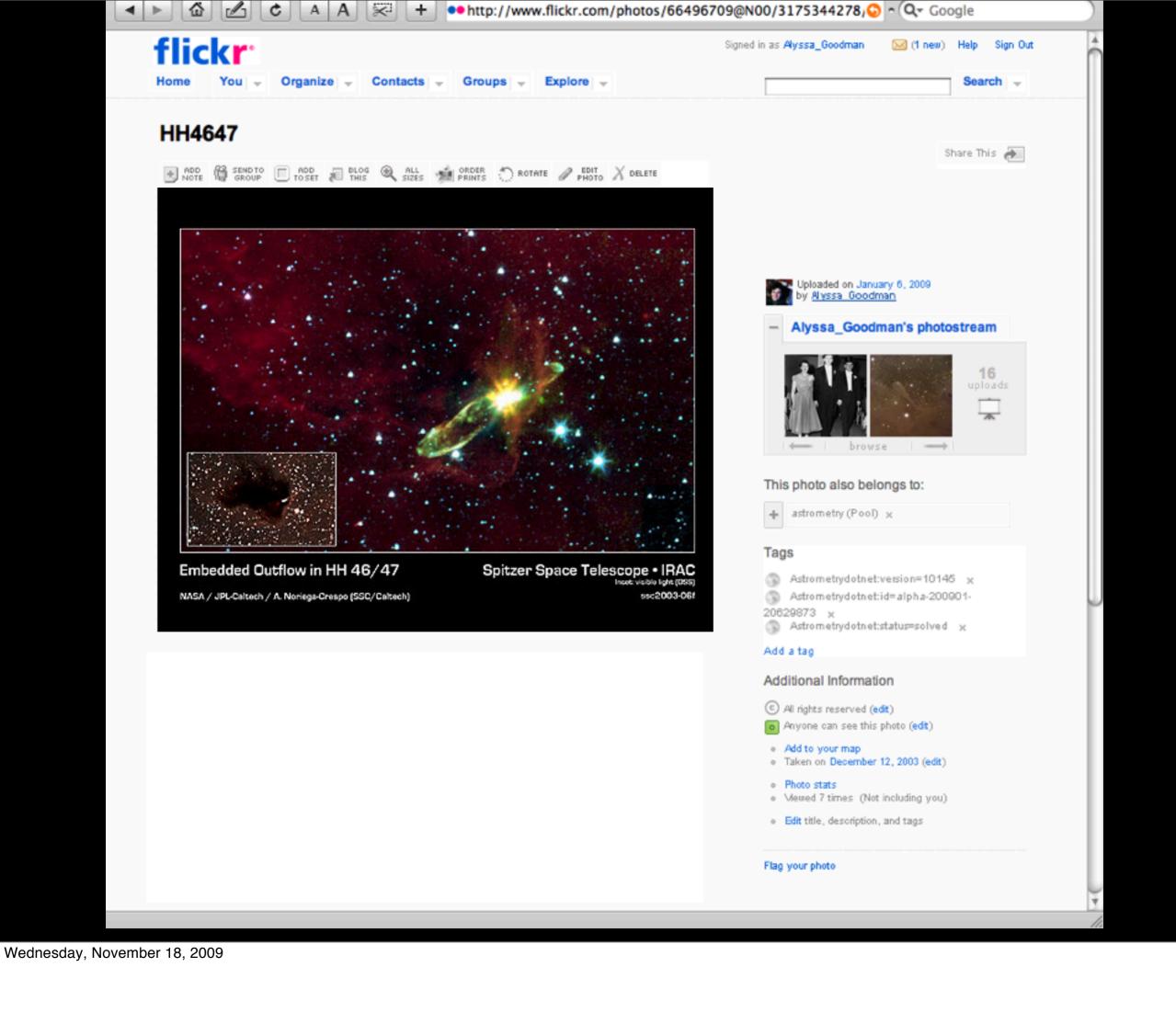
Dataset	Show	Perseus	Ophiuchus	Serpens	Link							
GBT: HI Data Cube	N	$\checkmark$	$\checkmark$	Ø	Data							
IRAS: Av/Temp Maps	N	$\checkmark$	$\checkmark$	√	Data							
FCRAO: 12CO	Y	⊻	⊻	$\checkmark$	Data							
FCRAO: 13CO		⊻	⊻	√	Data							
JCMT: 850 microns	N	⊻	⊻	Ø	Data							
Spitzer c2d: IRAC 1,3 (3.6,5.8 µm)	Y	$\checkmark$	$\checkmark$	$\checkmark$	Data							
Spitzer c2d: IRAC 2,4 (4.5,8 µm)	×	$\checkmark$	$\checkmark$	⊻	Data							
CSO/Bolocam: 1.2-mm	N	⊻	Ø	Ø	Data							
Spitzer MIPS: Derived Dust Map	×	⊻	Ø	Ø	Data							
Targeted Regions (Phase	e II, Sor	ne Data N	iot Yet Avail	able)								
CTIO/Calar Alto: NIR (J,H,Ks)	×	$\checkmark$	$\checkmark$	Ø	Data							
IRAM 30-m: N2H+ and C18O	Z	√	Ø	Ø	Data							
IRAM 30-m: 1.1-mm continuum	Z	$\checkmark$	Ø	Ø	Data							
Megacam/MMT: r,i,z images	V	$\checkmark$	Ø	Ø	Data							
Catalogs & Pointed Surveys												
NH3 Pointed Survey	V	$\checkmark$	Ø	Ø	Data							
YSO Candidate list (c2d)	¥	$\checkmark$	$\checkmark$	$\checkmark$	Data							

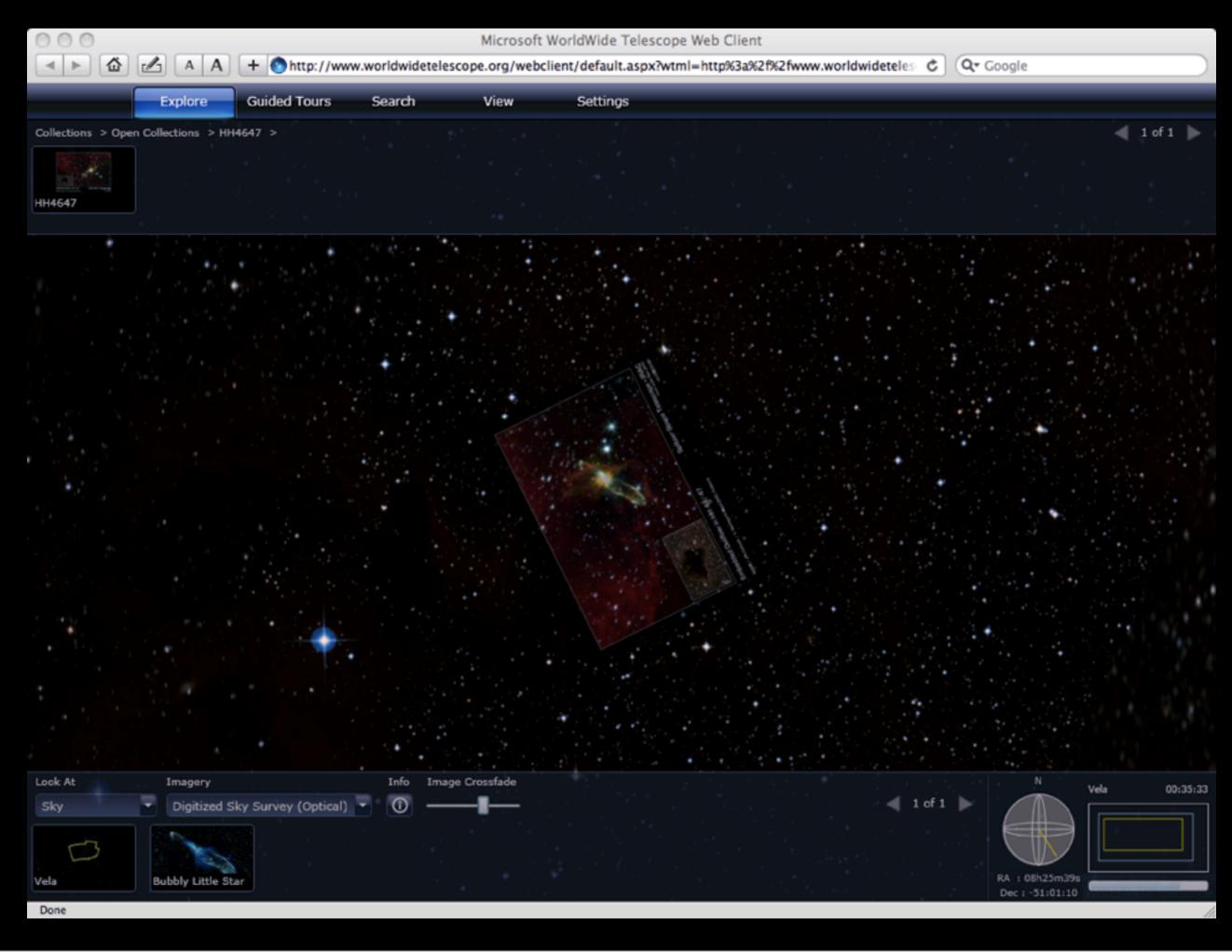
Done

000









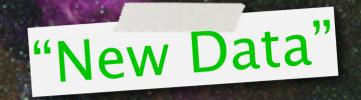
Wednesday, November 18, 2009







# "Old Data"



# "Your Data"

### 3D PDF

## "My Data"

**Context Search Filter** 

Digitized Sky Survey (Optical...

Imagery

Wednesday, November 18, 2009

0

mage Crossfade

## 3D PDF

#### LETTERS



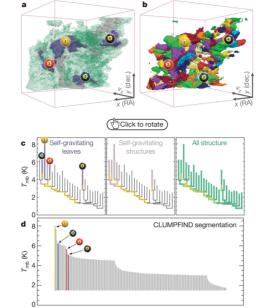


Figure 2 | Comparison of the 'dendrogram' and 'CLUMPFIND' featureidentification algorithms as applied to <sup>13</sup>CO emission from the L1448 region of Perseus. a, 3D visualization of the surfaces indicated by colours in the dendrogram shown in c. Purple illustrates the smallest scale selfgravitating 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-y) space. RA, right ascension; dec., declination. For comparison with the ability of dendrograms (c) to track hierarchical structure, **d** shows a pseudo-

dendrogram of the CLUMPFIND segmentation (**b**), with the same four labels used in Fig. 1 and in **a**. As 'clumps' are not allowed to belong to larger structures, each pseudo-branch in **d** is simply a series of lines connecting the maximum emission value in each clump to the threshold value. A very large number of clumps appears in **b** because of the sensitivity of CLUMPFIND to noise and small-scale structure in the data. In the online PDF version, the 3D cubes (**a** and **b**) can be rotated to any orientation, and surfaces can be turned on and off (interaction requires Adobe Acrobat version 7.0.8 or higher). In the printed version, the front face of each 3D cube (the 'home' view in the interactive online version) corresponds exactly to the patch of sky shown in Fig. 1, and velocity with respect to the Local Standard of Rest increases from front ( $-0.5 \text{ km s}^{-1}$ ) to back (8 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 two free parameters, the same molecular-line data set<sup>8</sup> can be used to show either that the frequency distribution of clump mass is the same as the initial mass function of stars or that it follows the much shallower mass function associated with large-scale molecular clouds (Supplementary Fig. 1).

Four years before the advent of CLUMPFIND, 'structure trees'<sup>9</sup> were proposed as a way to characterize clouds' hierarchical structure 64

©2009 Macmillan Publishers Limited. All rights reserved

### PDF



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 '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 (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  $M_{\text{lum}} = X_{13\text{CO}}L_{13\text{CO}}$  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,  $\alpha_{obs} = 5\sigma_{o}^{2}R/GM_{hm}$ 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 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 fields<sup>16</sup>, its measured value should only be used as a guide to the longevity (boundedness) of any particular feature.

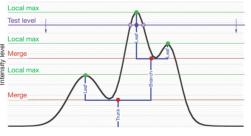
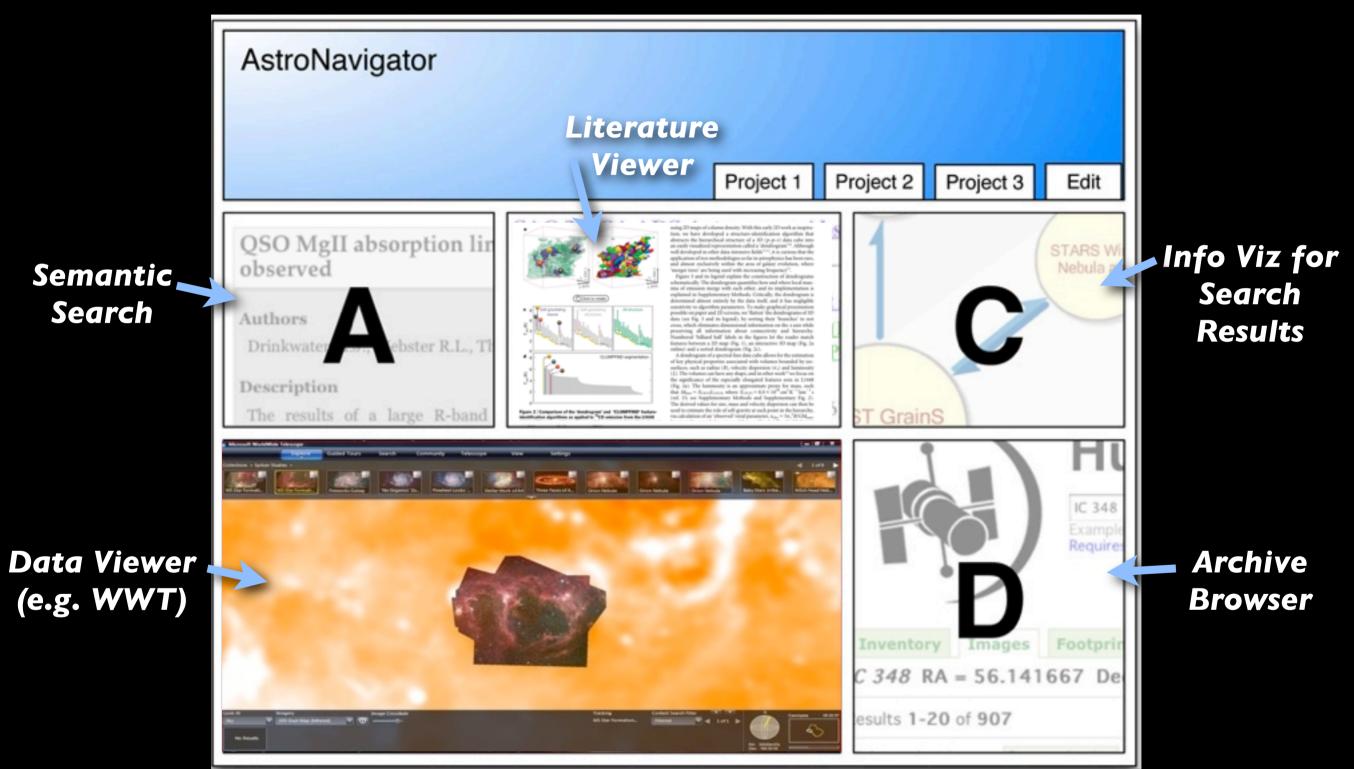


Figure 3 | Schematic illustration of the dendrogram process. Shown is the construction of a dendrogram from a hypothetical one-dimensional emission profile (black). The dendrogram (blue) can be constructed by 'dropping' a test constant emission level (purple) from above in tiny steps (exaggerated in size here, light lines) until all the local maxima and mergers are found, and connected as shown. The intersection of a test level with the emission is a set of points (for example the light purple dots) in one 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 manalogue 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.

# "Seamless Astronomy"



Mockup based on work of Eli Bressert, excerpted from NASA AISRP proposal by Goodman, Muench, Christian, Conti, Kurtz, Burke, Accomazzi, McGuinness, Hendler & Wong, 2008

### Seamless Astronomy

www.cfa.harvard.edu/~agoodman and worldwidetelescope.org

