

Visualization and Analysis of Rich Spectral Line Datasets

Elisabeth A.C. Mills (NRAO Jansky Fellow)

Jeff Kern, Brian Kent (NRAO), Joanna Corby (Uva)

Thanks to: E. Rosolowsky, A. Ginsburg

New Radio Telescopes

ALMA

66 dishes
wavelengths: 0.3 - 3.6 mm

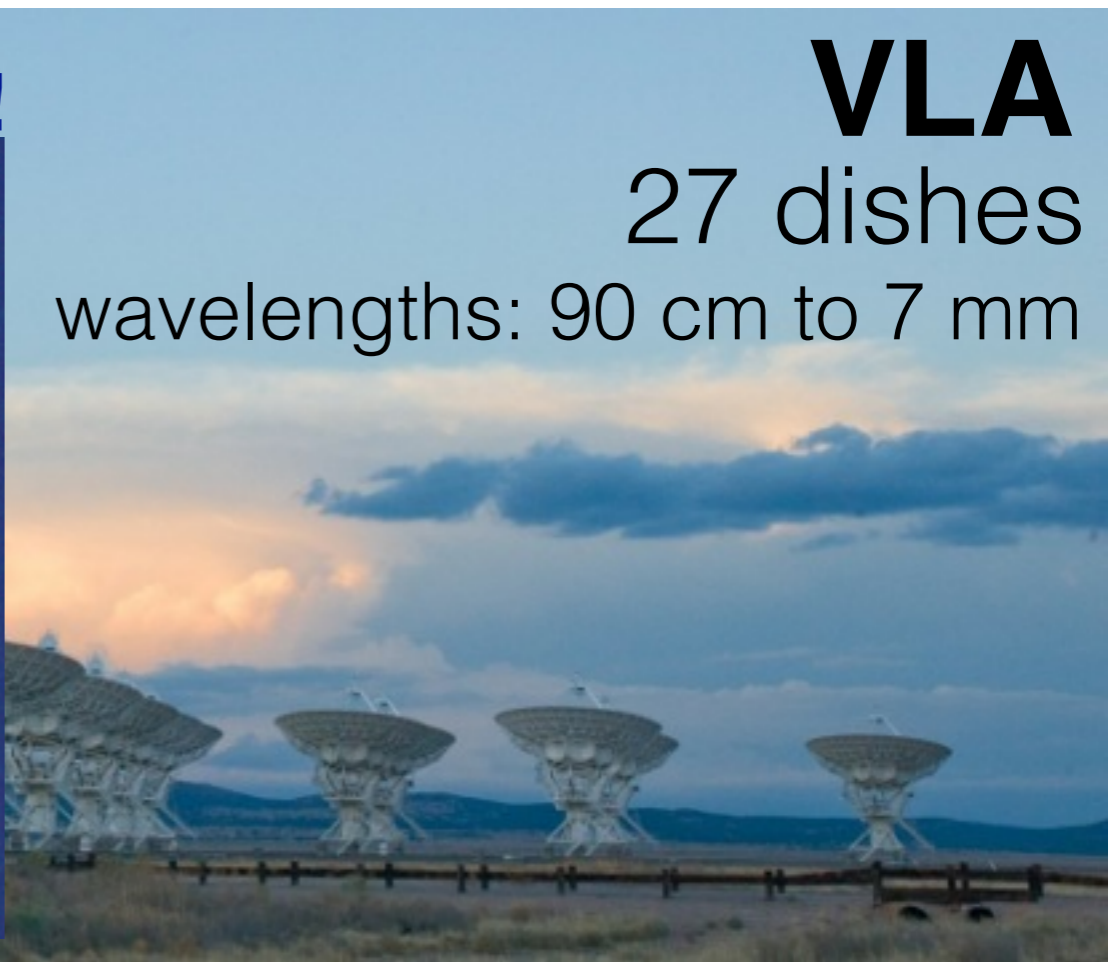


it's new on the inside!

Parameter	SPECIFICATIONS	
	VLA	EVLA
Continuum sensitivity in 12 hrs. 1 σ	10 μ Jy	0.8 μ Jy
Maximum bandwidth	0.1 GHz	8 GHz
Number of frequency channels at maximum bandwidth	16	16,384
Maximum number of frequency channels	512	4,194,304
(Log) Frequency coverage, 1 - 50 GHz	22%	100%
Number of baselines	351	351
Spatial Resolution (5GHz)	0.3 arcsec	0.3 arcsec

VLA

27 dishes
wavelengths: 90 cm to 7 mm



The Big Data Problem

**More telescope dishes = more sensitivity/
resolution**

**More bandwidth = more sensitivity, more
spectral lines**

**More channels = better spectral resolution,
RFI excision**

**Faster integration = increased ability to do
time-domain astronomy**

... and more data.

Big Data Challenge #1: Volume

At full operations, ALMA will produce > 250 TB / year

VLA will soon do an all-sky survey, ~ 550 TB over ~ 5 years

This data volume is challenging to:

- store
- provide to users
- calibrate
- image

It could be worse:

ALMA capability:

250 TB/ **Day**

VLA capability:

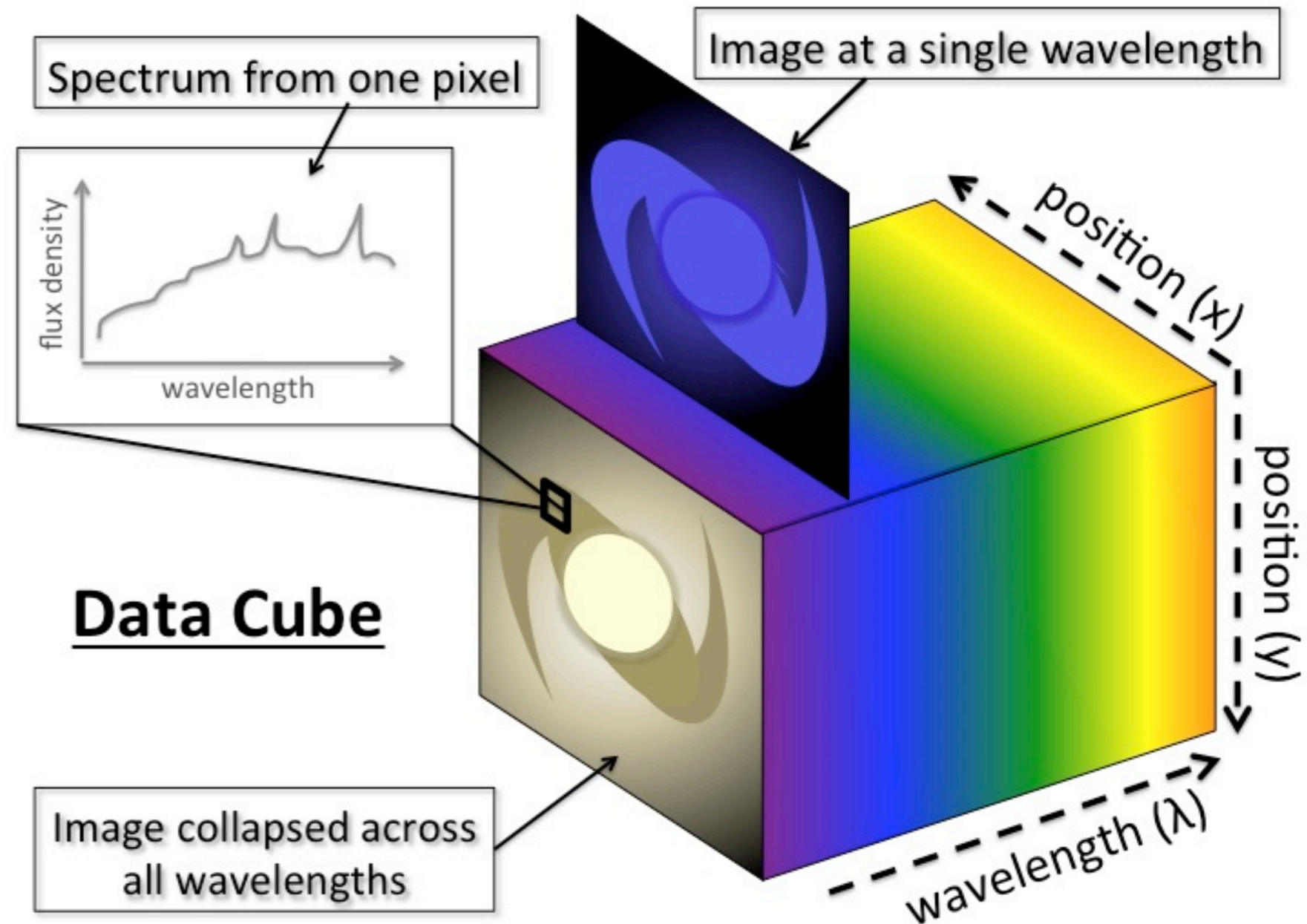
1.5 PB / **Day**



“...our pipeline is people”

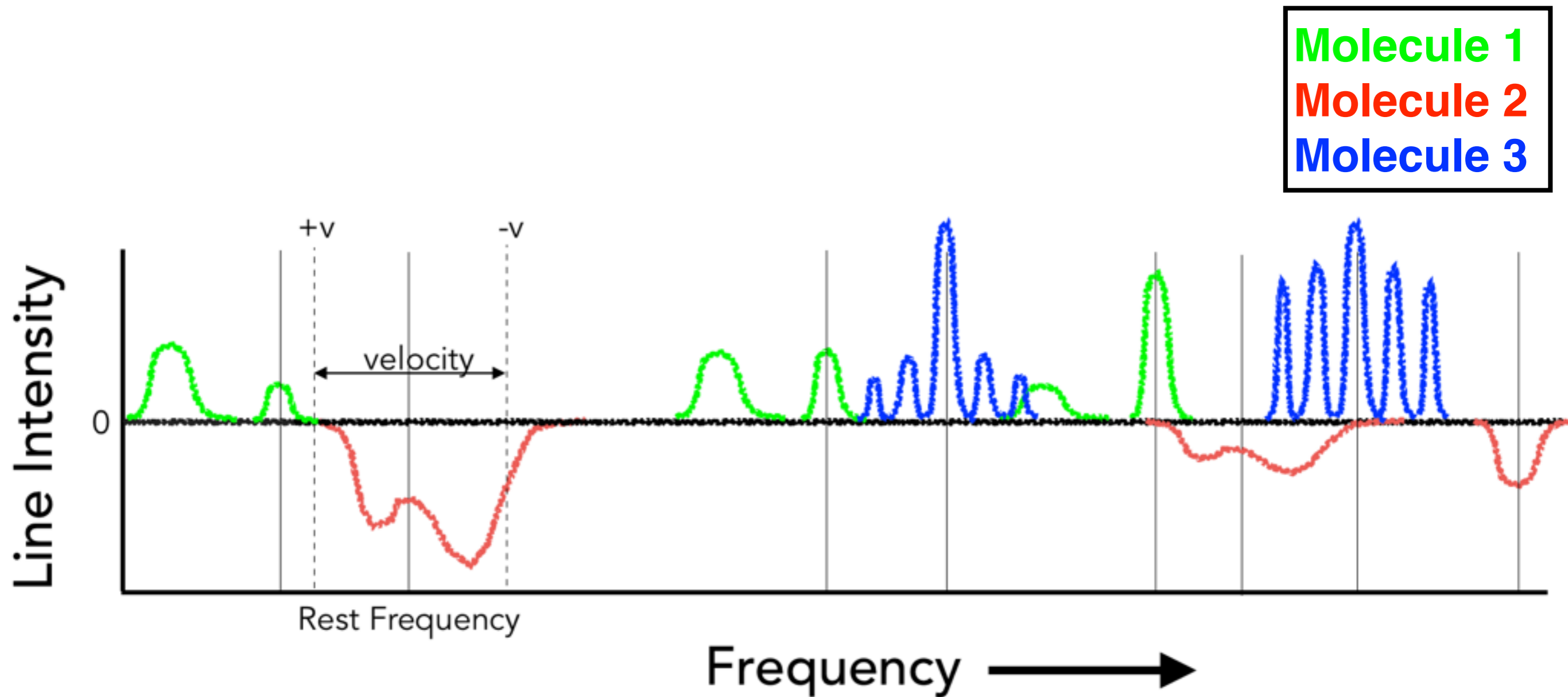
Big Data Challenge #2: Dimension

Spectral cubes are high-dimensional data sets



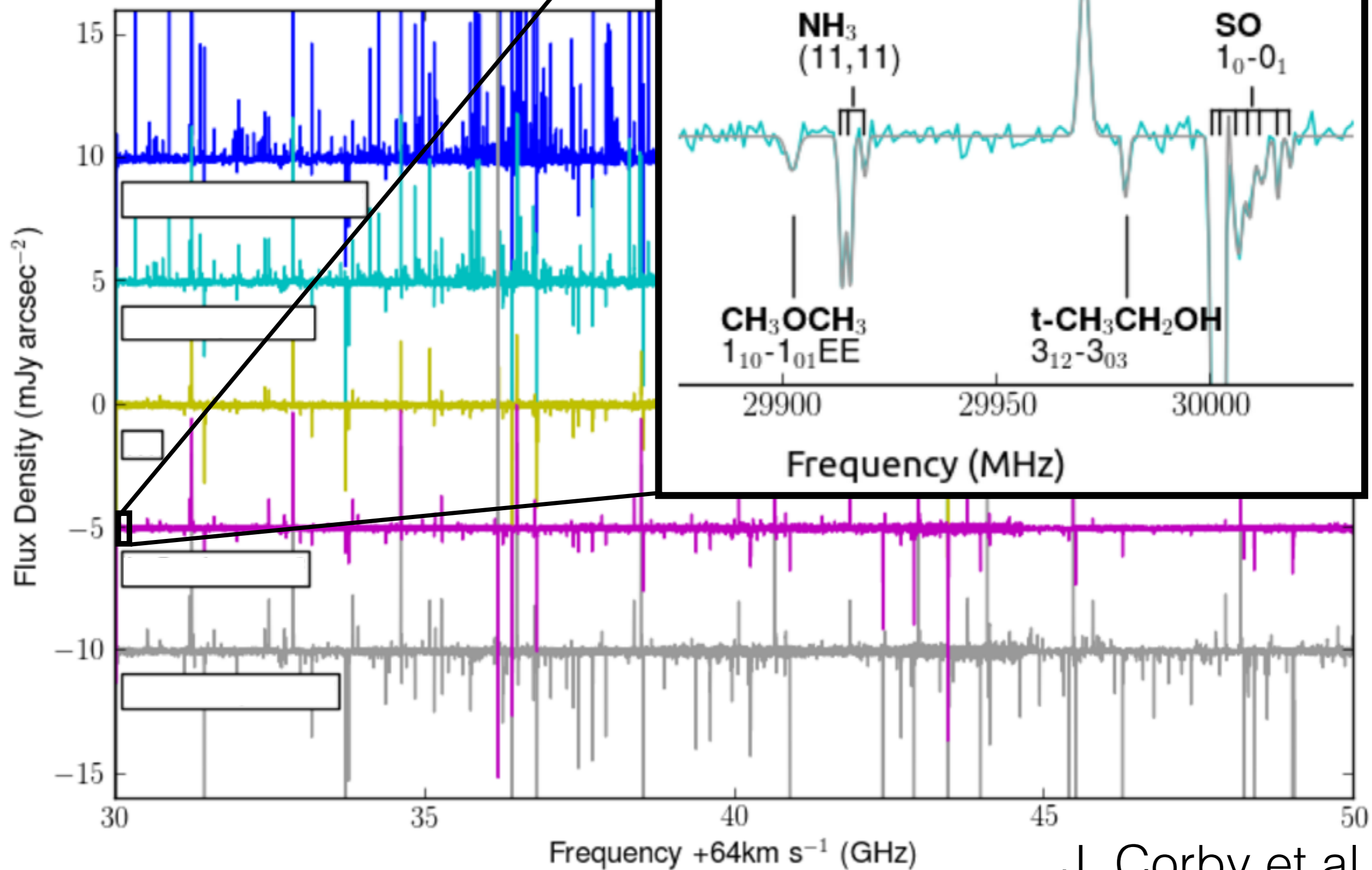
VLA: up to 4 million frequency channels

Big Data Challenge #3: Complexity



Ultimately one can observe lines of many different molecules with different structure at different velocities

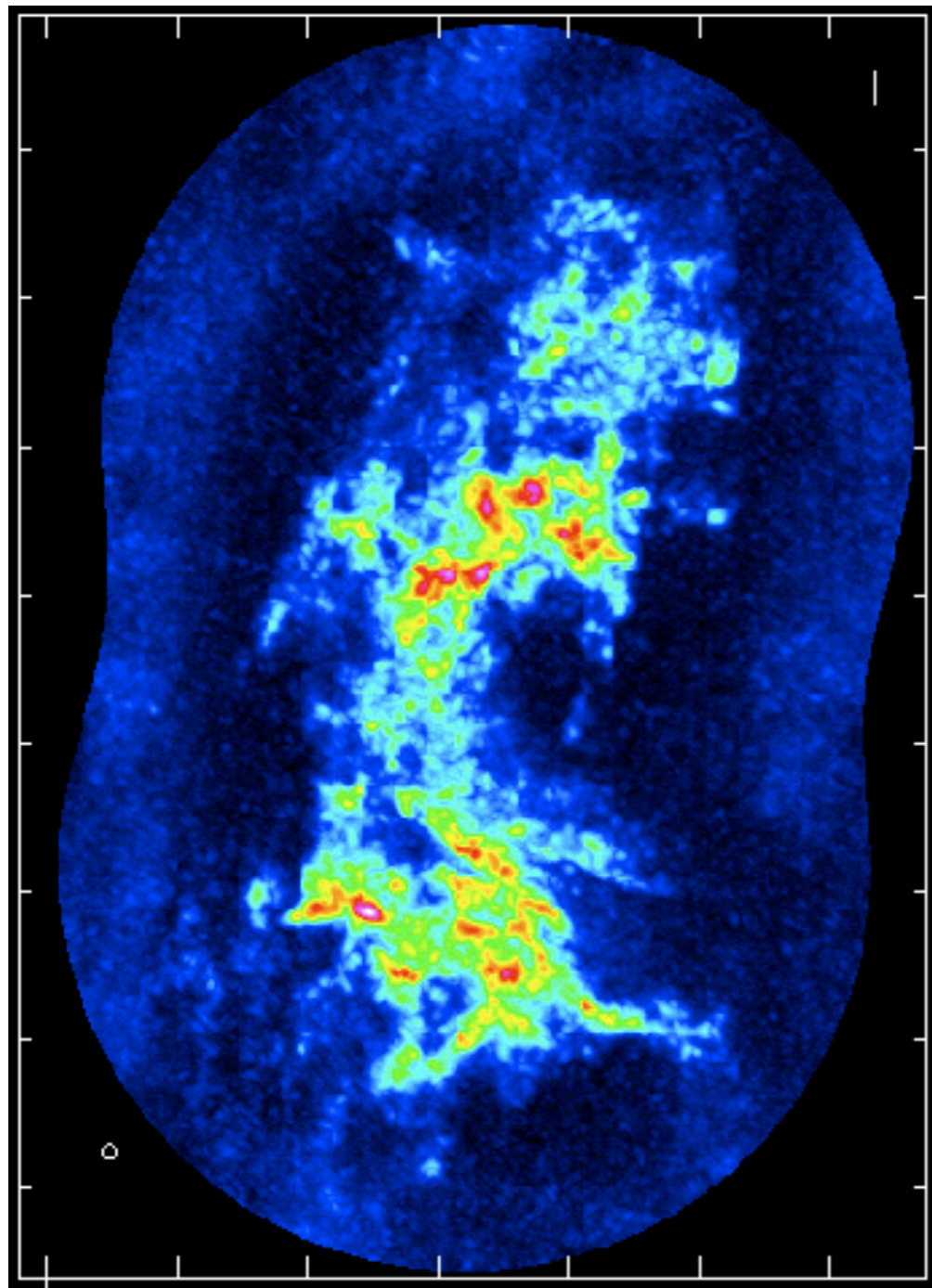
Real Spectra



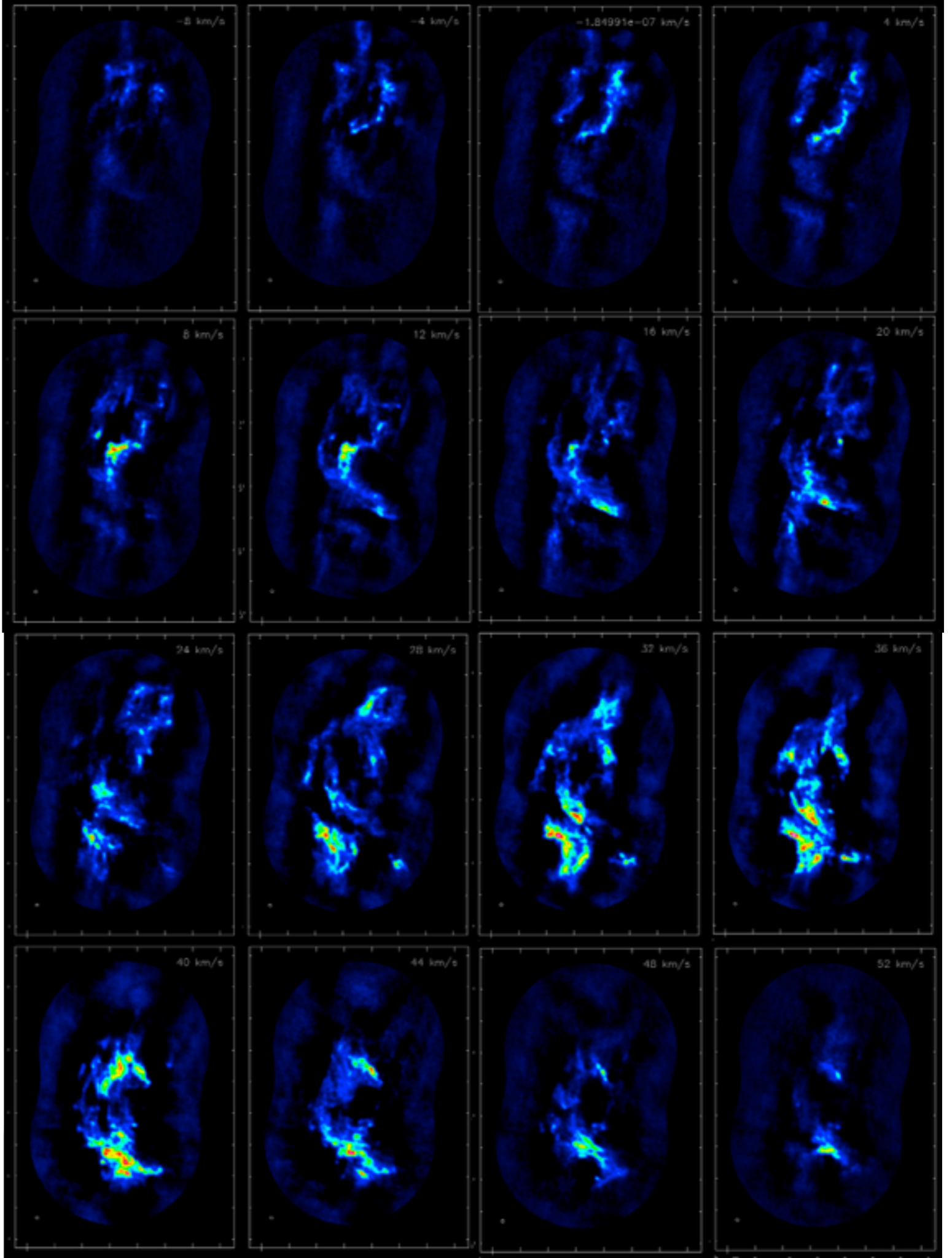
J. Corby et al.

Real Cubes

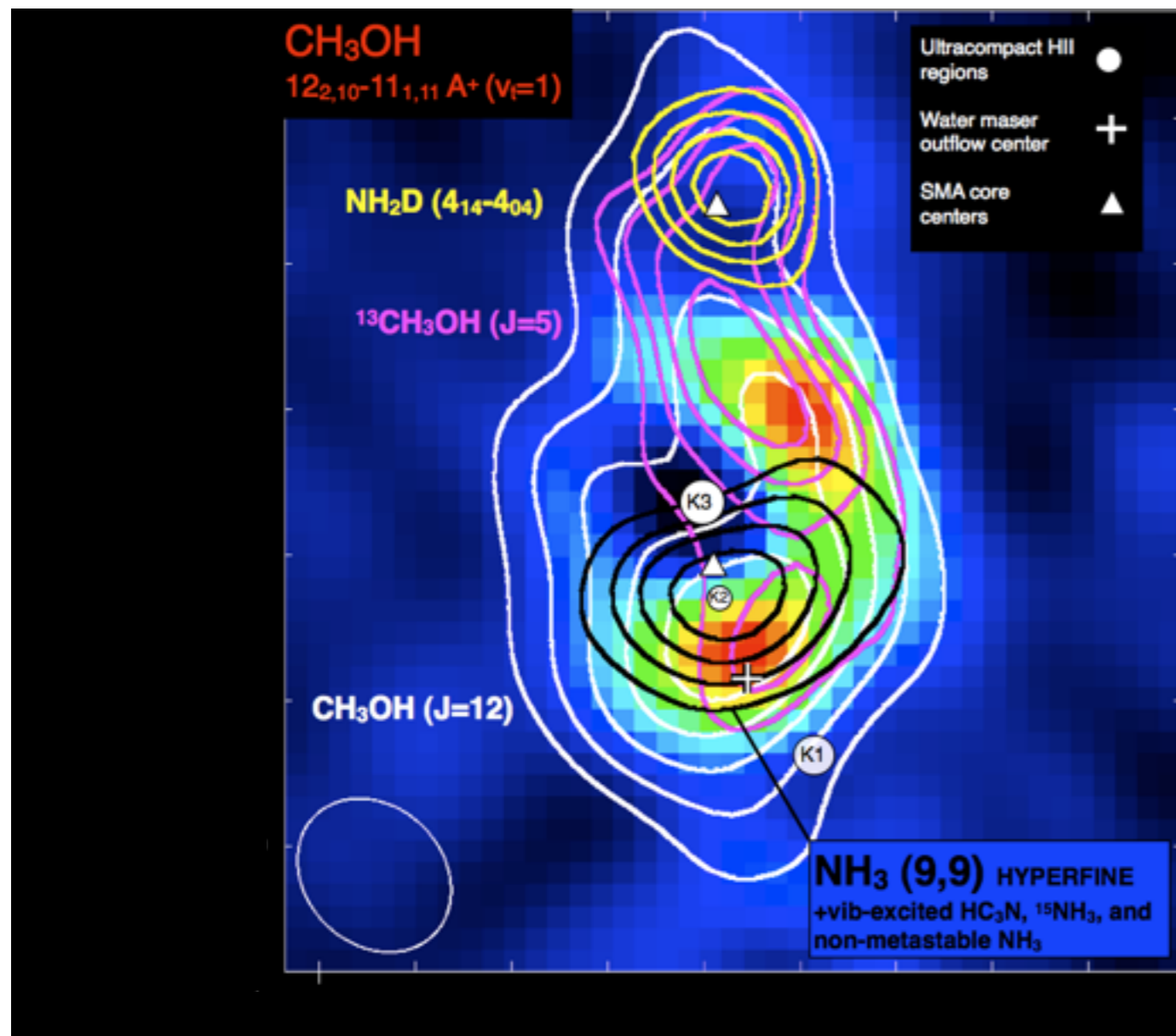
a complex spectrum at every position in this complex image.



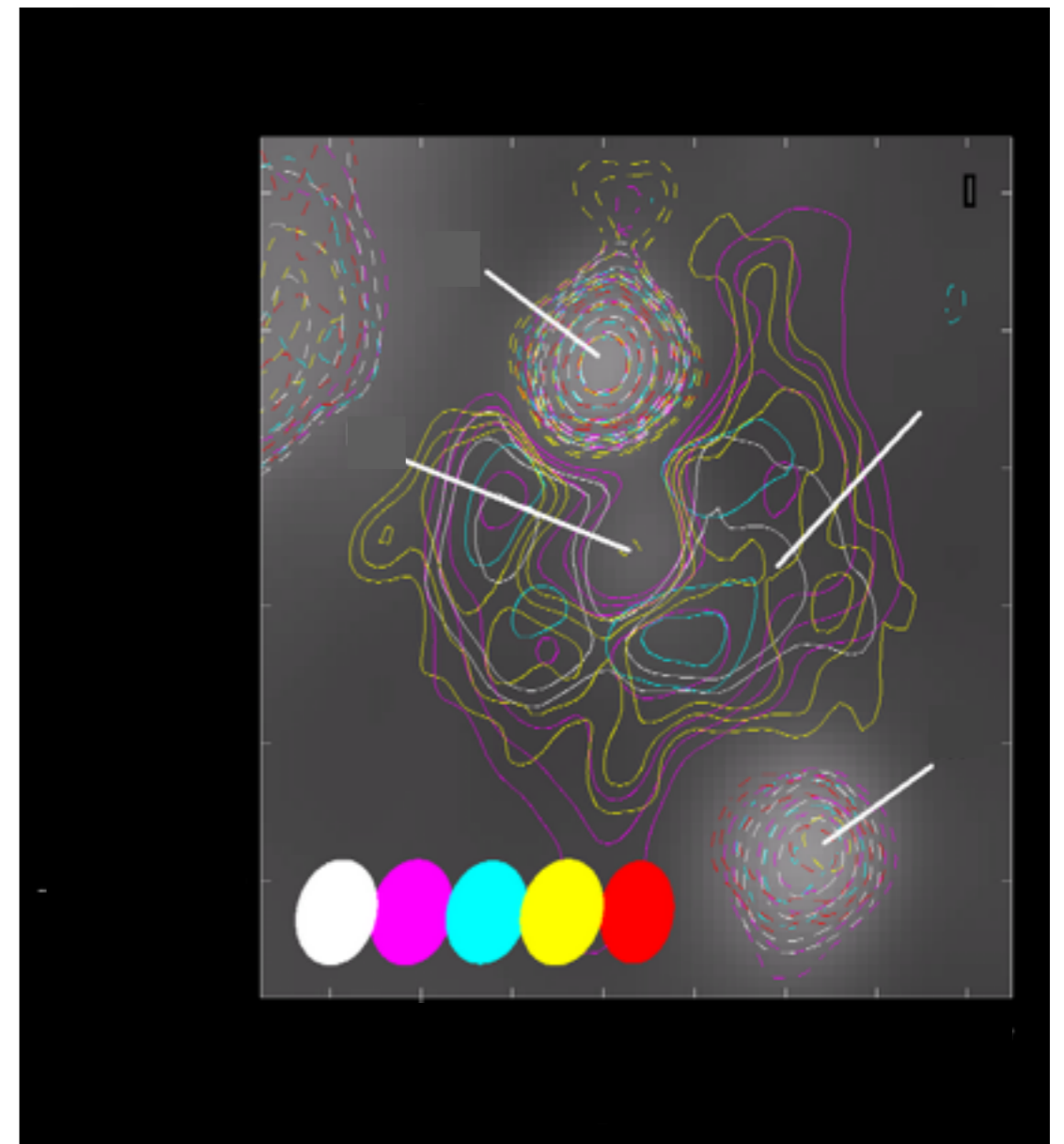
(E.A.C. Mills)



Balance losing & visualizing info?



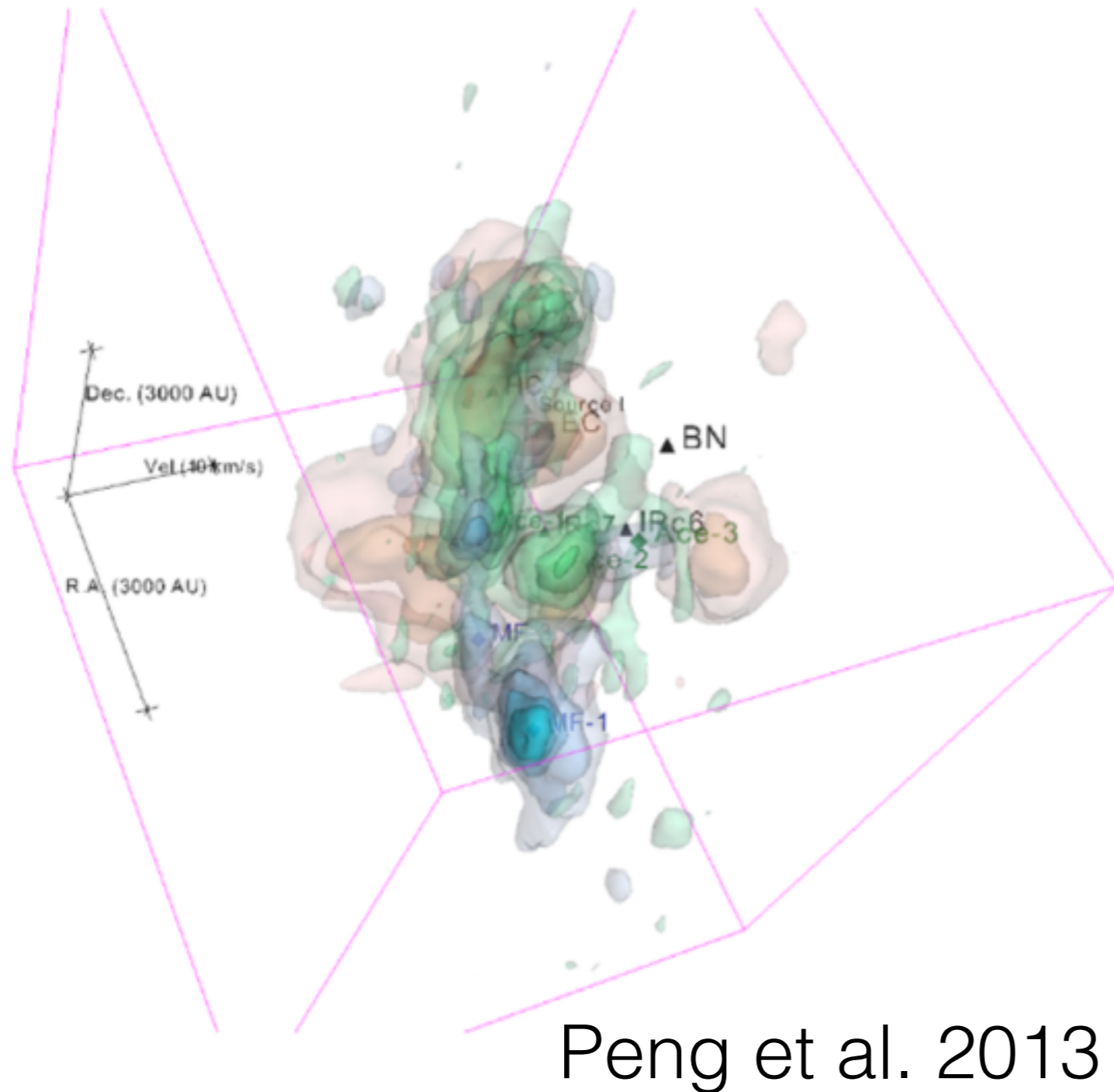
(E.A.C. Mills)



(J. Corby)

“lose-lose” plots of Integrated emission from molecules—
losing the velocity information AND really can’t quickly
navigate this information

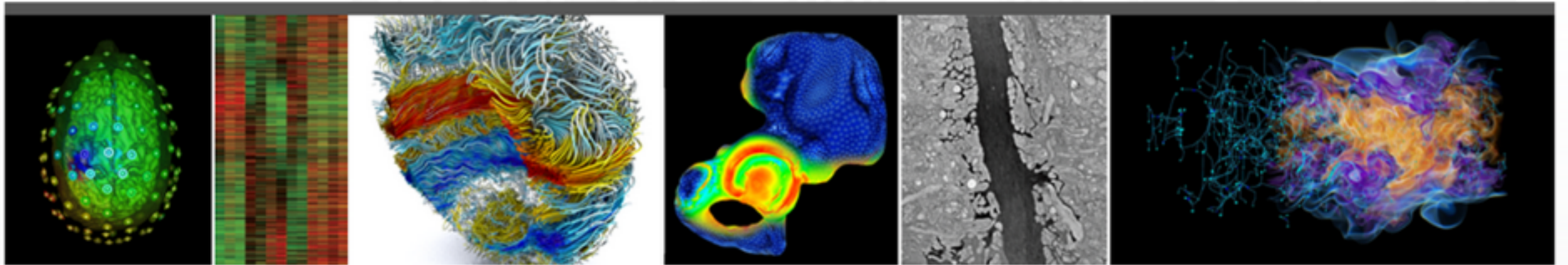
Can we stay tied to spatial representations?



3D rendering can provide an intuitive comparison, but still becomes complicated when more than a few species are compared.

Scientific Computing & Imaging Institute

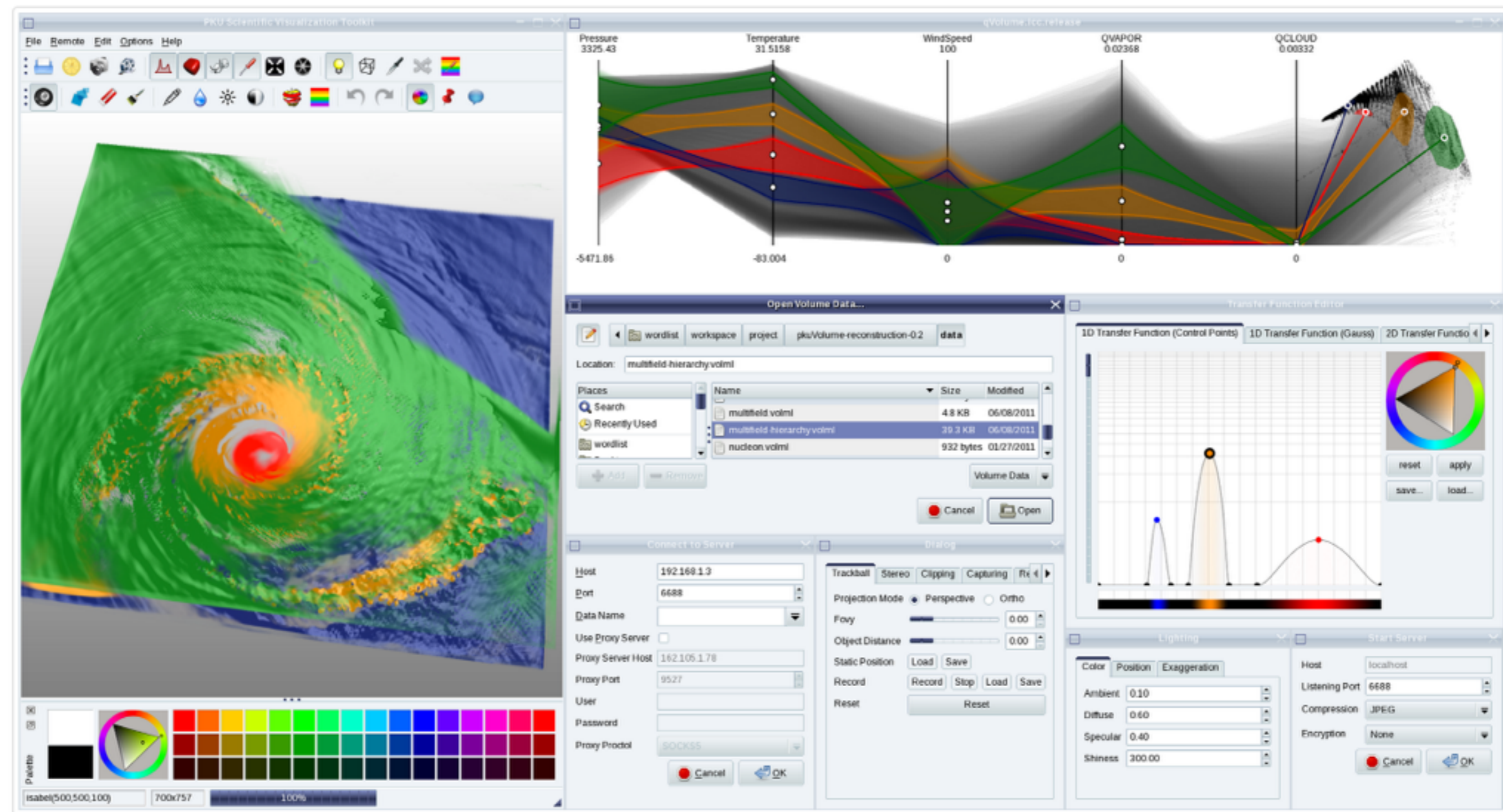
Calling in outside help!



- Scientific Visualization
- Information Visualization
- Uncertainty Visualization
- Volume Rendering
- Segmentation
- Topological Methods

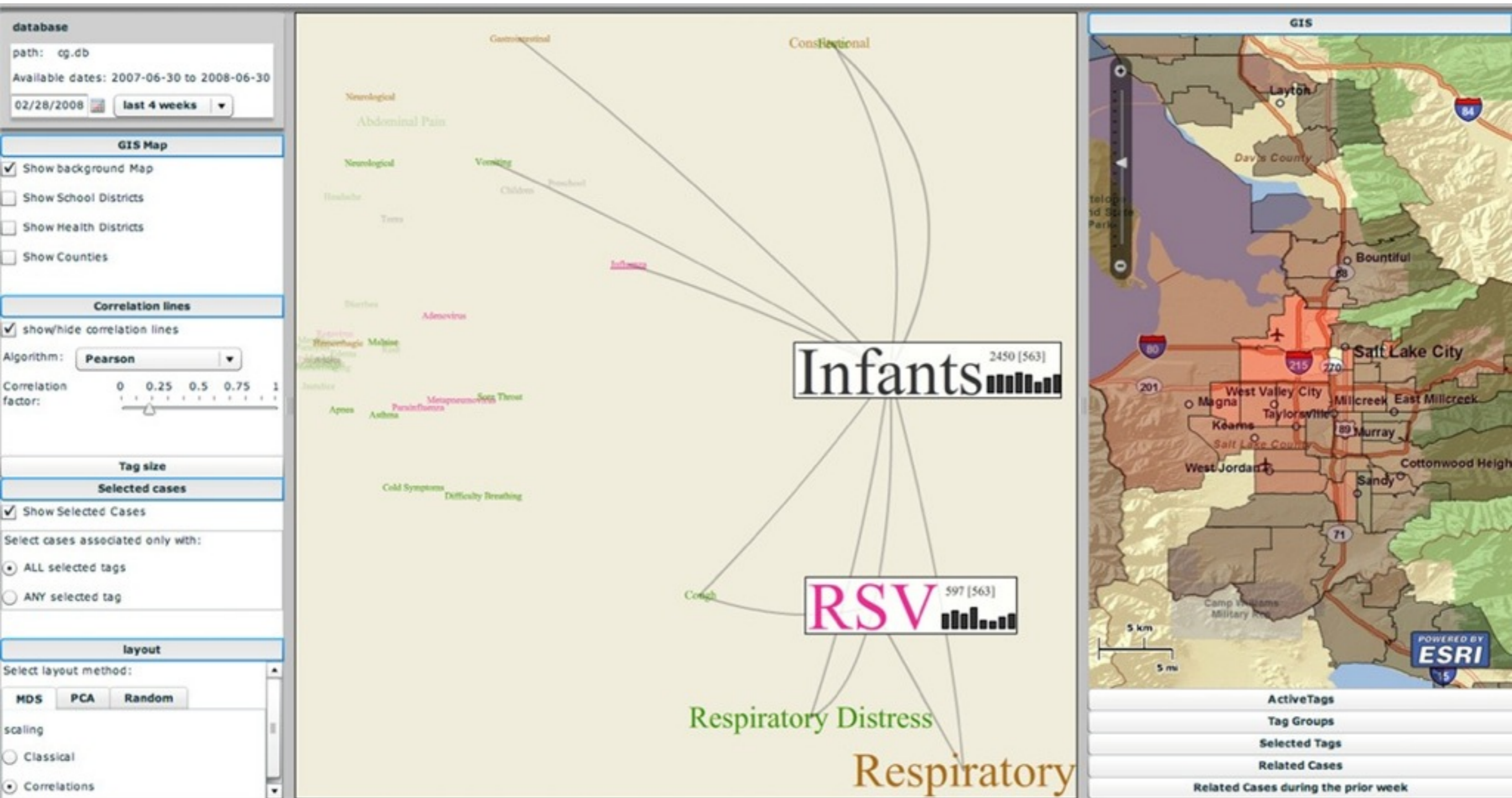
Interactively exploring many data cubes

Multidimensional Transfer Functions & Linked Views



(Image: PKU Visualization Toolkit, Peking University)

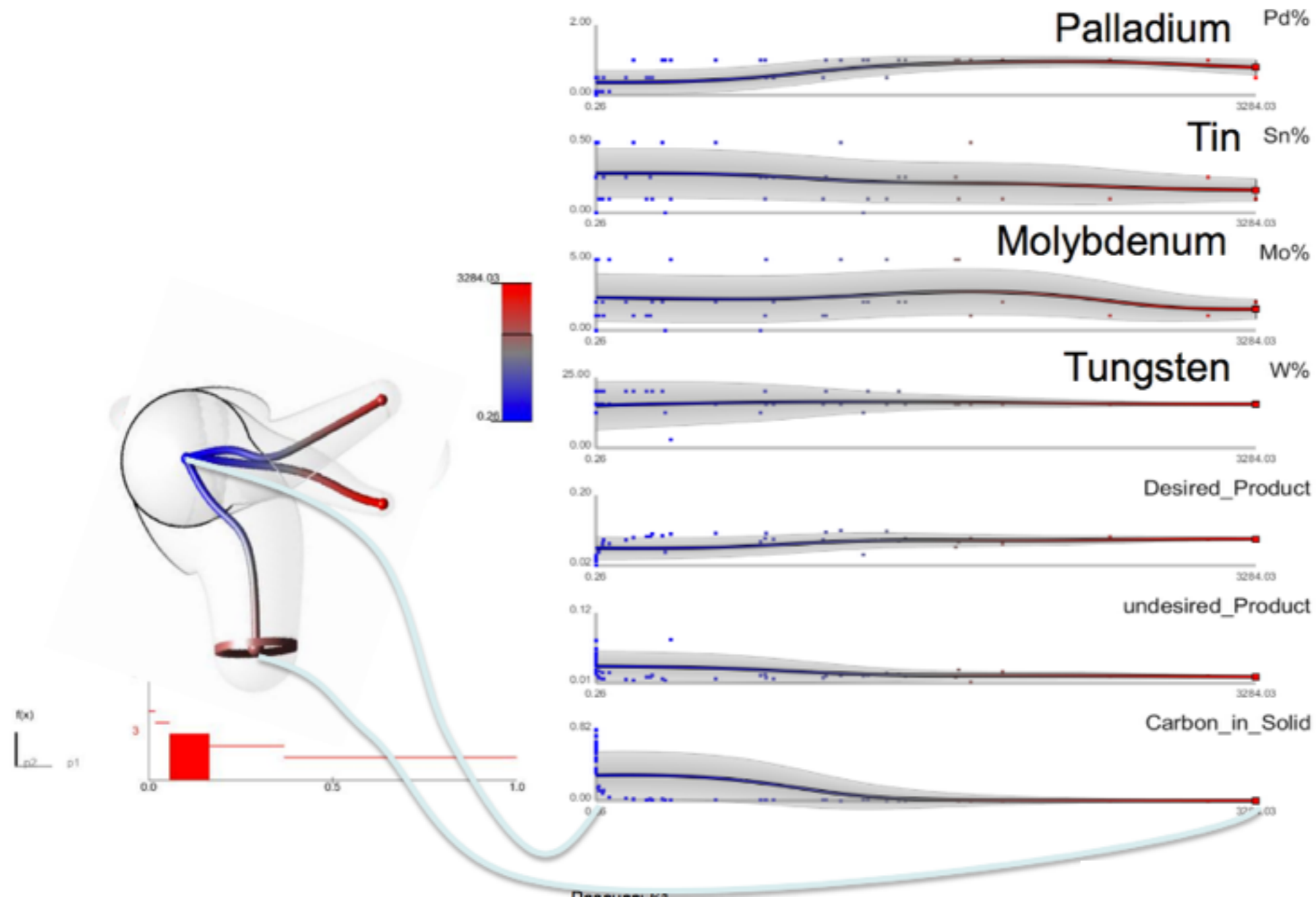
Metadata-aided Exploration



EpiCanvas: A. Yivnat (SCI)

Ignoring spatial information

Topological analysis



I. Pascucci (SCI)

3 Takeaways

- Look up SCl Institute: <https://www.sci.utah.edu/>
- Ask me about synergy with the VLA Sky Survey
- Talk to me about dealing with complex datasets