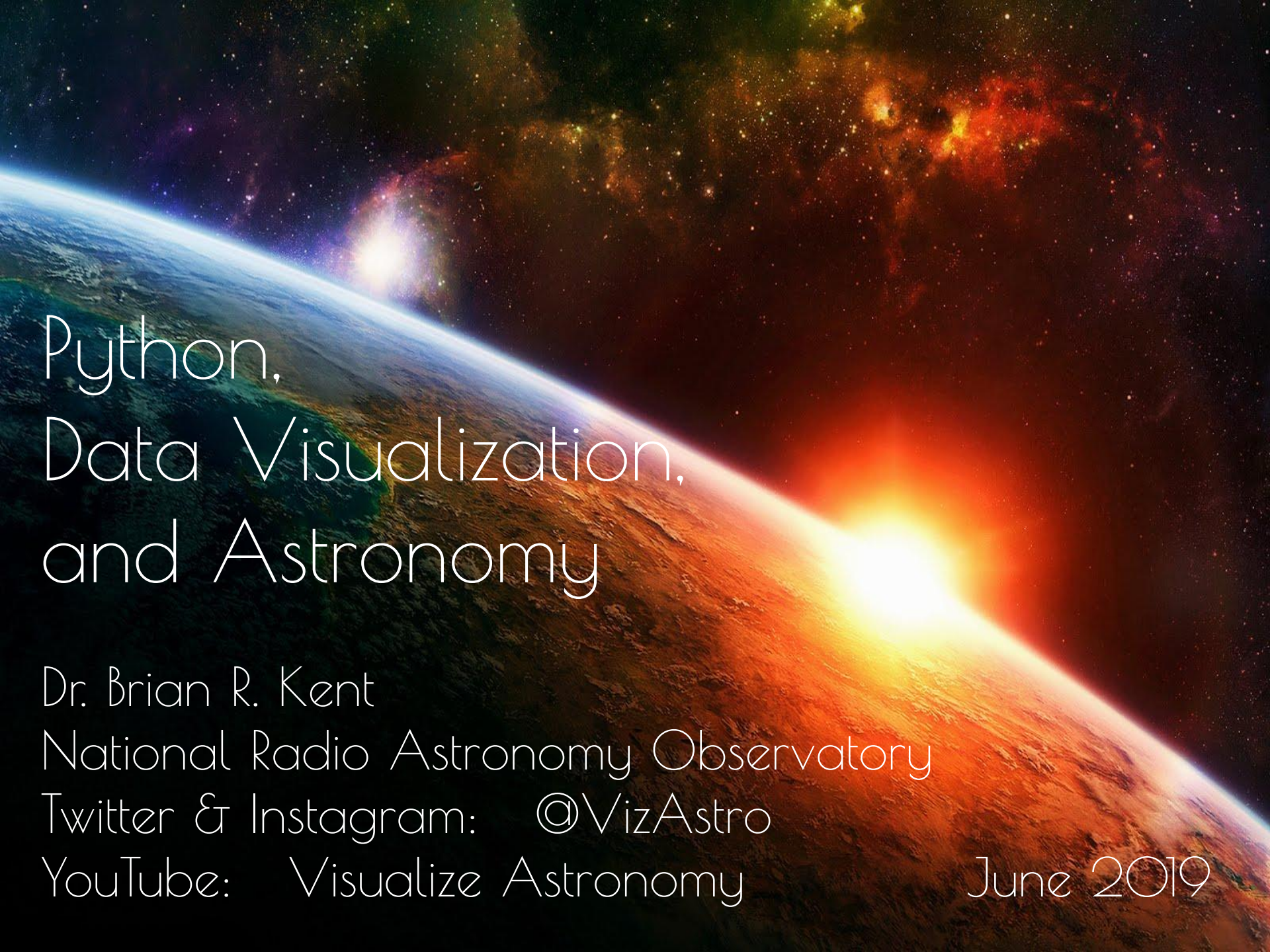




<https://www.youtube.com/watch?v=n04M39SbfW0>



# Python, Data Visualization, and Astronomy

Dr. Brian R. Kent

National Radio Astronomy Observatory

Twitter & Instagram: @VizAstro

YouTube: Visualize Astronomy

June 2019

# Overview

- History of Python
- Why use Python in Astronomy
- Short Introduction to Revision Control
- Google Colab
- Some of my research - 3D rendering of Astronomical Data
- Resources to explore further
- Conferences (in-person and remote!)



# Python

- High level programming language that is scriptable and supports many programming styles.
- Created in the late 80s/early 90s by Guido van Rossum, named after *Monty Python*
- The **Python Standard Library** can be used to accomplish a wide variety of tasks - always check there first!



# tiobe.com - Usage of Languages

Jun 2018	Jun 2017	Change	Programming Language	Ratings	Change
1	1		Java	15.368%	+0.88%
2	2		C	14.936%	+8.09%
3	3		C++	8.337%	+2.61%
4	4		Python	5.761%	+1.43%
5	5		C#	4.314%	+0.78%
6	6		Visual Basic .NET	3.762%	+0.65%
7	8	▲	PHP	2.881%	+0.11%
8	7	▼	JavaScript	2.495%	-0.53%
9	-	▲▲	SQL	2.339%	+2.34%
10	14	▲▲	R	1.452%	-0.70%
11	11		Ruby	1.253%	-0.97%
12	18	▲▲	Objective-C	1.181%	-0.78%
13	16	▲	Visual Basic	1.154%	-0.86%
14	9	▼▼	Perl	1.147%	-1.16%
15	12	▼	Swift	1.145%	-1.06%
16	10	▼▼	Assembly language	0.915%	-1.34%
17	17		MATLAB	0.894%	-1.10%
18	15	▼	Go	0.879%	-1.17%
19	13	▼▼	Delphi/Object Pascal	0.875%	-1.28%
20	20		PL/SQL	0.848%	-0.72%

# TIOBE Index for June 2019



## June Headline: Python continues to soar in the TIOBE index

This month Python has reached again an all time high in TIOBE index of 8.5%. If Python can keep this pace, it will probably replace C and Java in 3 to 4 years time, thus becoming the most popular programming language of the world. The main reason for this is that software engineering is booming. It attracts lots of newcomers to the field. Java's way of programming is too verbose for beginners. In order to fully understand and run a simple program such as "hello world" in Java you need to have knowledge of classes, static methods and packages. In C this is a bit easier, but then you will be hit in the face with explicit memory management. In Python this is just a one-liner. Enough said.

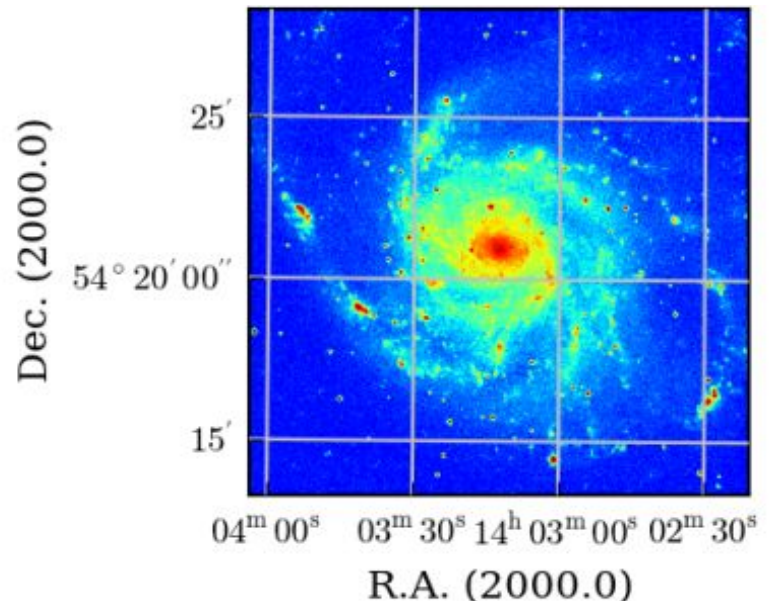
The TIOBE Programming Community index is an indicator of the popularity of programming languages. The index is updated once a month. The ratings are based on the number of skilled engineers world-wide, courses and third party vendors. Popular search engines such as Google, Bing, Yahoo!, Wikipedia, Amazon, YouTube and Baidu are used to calculate the ratings. It is important to note that the TIOBE index is not about the *best* programming language or the language in which *most lines of code* have been written.

The index can be used to check whether your programming skills are still up to date or to make a strategic decision about what programming language should be adopted when starting to build a new software system. The definition of the TIOBE index can be found [here](#).

Jun 2019	Jun 2018	Change	Programming Language	Ratings	Change
1	1		Java	15.004%	-0.36%
2	2		C	13.300%	-1.64%
3	4	▲	Python	8.530%	+2.77%
4	3	▼	C++	7.384%	-0.95%
5	6	▲	Visual Basic .NET	4.624%	+0.86%
6	5	▼	C#	4.483%	+0.17%
7	8	▲	JavaScript	2.716%	+0.22%
8	7	▼	PHP	2.567%	-0.31%
9	9		SQL	2.224%	-0.12%
10	16	▲▲	Assembly language	1.479%	+0.56%
11	15	▲▲	Swift	1.419%	+0.27%
12	12		Objective-C	1.391%	+0.21%
13	11	▼	Ruby	1.388%	+0.13%
14	60	▲▲	Groovy	1.300%	+1.11%
15	18	▲	Go	1.257%	+0.38%
16	14	▼	Perl	1.173%	+0.03%
17	19	▲	Delphi/Object Pascal	1.129%	+0.25%
18	17	▼	MATLAB	1.077%	+0.18%
19	13	▼▼	Visual Basic	1.069%	-0.08%
20	20		PL/SQL	0.929%	+0.08%

# Python and Astronomy

- With iPython, it is used as the command line shell & interpreter for CASA
- Used for the framework for VLA/GBT imaging/ALMA pipelines
- Can interface with AIPS via Obit
- Can interface with IRAF via PyRAF
- Used in PRESTO - pulsar data reduction
- Many modules and libraries available - numpy, matplotlib, Kapteyn, astropy, ApIPy, etc.



# Popular Python Resources

**iPython** - more user friendly shell

**Astropy** - great for data import, manipulation,  
catalog queries

**Matplotlib** - general purpose plotting tool

**Scipy** - numpy and fitting routines (some  
overlap with astropy...)

**Kapteyn** - Good for mapping projections

**AstroML** - Machine Learning

<http://www.astroml.org/index.html>

**PANDAS** - Data Analysis Library

<http://pandas.pydata.org/>



# Where to start?

- <https://safe.nrao.edu/wiki/bin/view/Main/PythonResources>
- <http://www.astropython.org/resources>
- Pycon 2019: <https://us.pycon.org/2019/>



**Friendly VRI: Control Window**

File Help

North-South (km)

East-West (km)

Telescope	Array
PdB	B
PdB	C
PdB	D
SMA	Compact
SMA	Compact_N
SMA	Extended
SMA	Subcompact
SMA	VExtended
VLA	A
VLA	B
VLA	C
VLA	D
WSRT	A

Hour Angle Range (hours):

-12.0 -6.0 0.0 6.0 12.0

-2.03 2.03

Sampling Cadence (s): 300

Add

Telescope	Array	HA-Start	HA-End	Cadence
VLA	B	-2.03	2.03	300

Clear Selected    Clear All

Plot Elevation

Antenna  $\theta$ : 25.0 m    Min Baseline: 0.243 km

Telescope  $\phi$ : 34.0787°    Max Baseline: 11.128 km

Model Image:     Source Declination (degrees): 20.0    Observing Frequency (MHz): 1420.0

Pixel Scale (arcsec): 1.0    864 x 864 px / 14.40' x 14.40'    Robust Weighting Factor: None

Array

Model

INPUTS

uv-Coverage

Calculate & Plot

Model FFT

Calculate & Plot

uv-Grid

Beam

Do Observation

Observation

Show Plot Window

**Friendly VRI: Plot Window**

Model Image

uv-Coverage

Synthesised Beam

Model FFT

Observed FFT

Observed Image

Properties of uv-Coverage

Scale Range: 3.91° to 3.27°

uv-Range: 1.051 to 52.709 k $\lambda$

Properties of Model FFT:

FFT Pixel Scale: 0.477 x 0.477 k $\lambda$

FFT Image Limit: 206.265 k $\lambda$

Show Control Window

<https://crpurcell.github.io/friendlyVRI/>

# Example: Create a noisy spectrum and fit a Gaussian

```
#!/pip install astropy

import pylab
import os
import sys
from math import sqrt
import numpy as np
import matplotlib.pyplot as plt
from astropy.modeling import models, fitting
```

# Google Colab

Colaboratory is a **Google** research project created to help disseminate machine learning education and research. It's a Jupyter notebook environment that requires no setup to use and runs entirely in the cloud.

<https://colab.research.google.com/drive/1zrYZMjcNBUmYn2xbejeZoSmlu-JCbofJ>

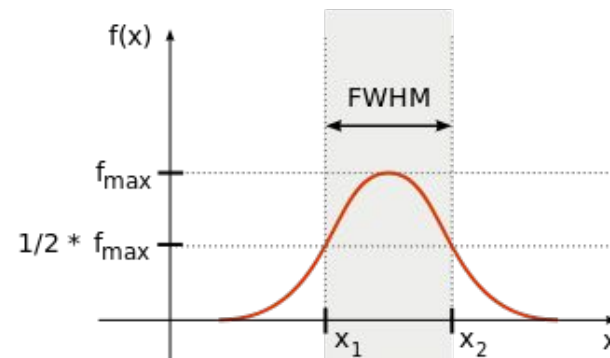
# Functions to Define

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-x_0)^2}{2\sigma^2}\right]$$

Usage:

```
def peval(x, p):  
    # The model function with parameters p  
    return (1./sqrt(2*np.pi*p[1]**2))*np.exp(-(x-p[0])**2/(2*p[1]**2))  
  
# Generate model data for a Gaussian with param mu and sigma and add noise  
x = np.arange(-10.,10.,20./1000) #1000 points  
preal = [-2, .5]  
y_true = peval(x,preal)  
mu,sigma = 0,0.7  
  
#Create my spectrum  
y = y_true + 0.06 * np.random.normal(mu,sigma, len(x) )
```

$$\text{FWHM} = 2\sqrt{2\ln 2} \sigma \approx 2.355 \sigma.$$





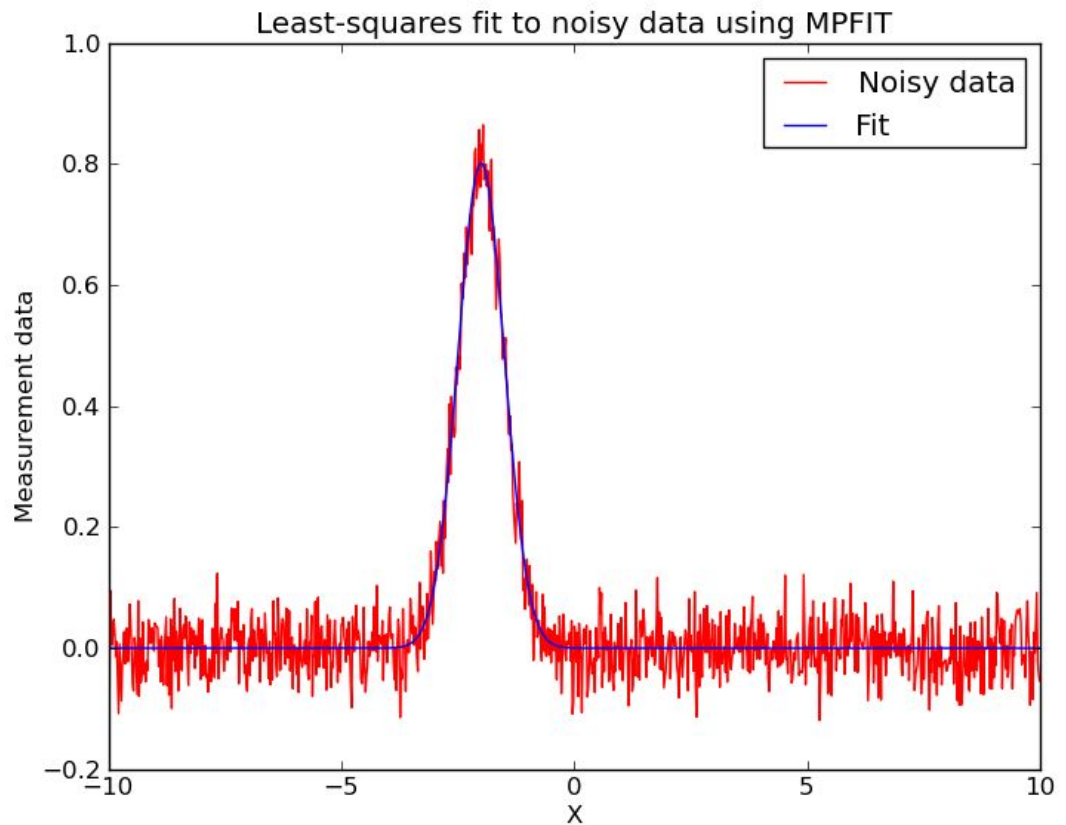
## Implementation of Levenberg-Marquardt least-squares minimization

```
# Fit the data using a Gaussian
g_init = models.Gaussian1D(amplitude=0.7, mean=-2, stddev=1.)
fit_g = fitting.LevMarLSQFitter()
g = fit_g(g_init, x, y)
```

$$S(\boldsymbol{\beta}) = \sum_{i=1}^m [y_i - f(x_i, \boldsymbol{\beta})]^2$$

# Execution

```
# Plot the data with the best-fit model  
plt.figure(figsize=(16,10))  
plt.plot(x, y, 'r', label='Noisy Data')  
plt.plot(x, g(x), label='Gaussian')  
plt.xlabel('X')  
plt.ylabel('T')  
plt.legend(loc=2)  
plt.show()
```



# Revision Control

# Revision Control for Software

Popular revision control systems (RCS) include SVN, CVS, Mercurial, and git.

Tracks changes made to files - good for collaboration among teams or to see when and where changes were made to your code!

Let's do an example with git...

<https://bitbucket.org/brkent/demo>

# git revision control

```
mkdir sumstudent
```

```
cd sumstudent
```

```
git clone
```

```
https://brkent@bitbucket.org/brkent/demo.git
```

This will create and link your directory to the git repository



# git revision control

#Create a readme file

```
touch README
```

```
git add README
```

```
git commit -m "Initial README commit"
```

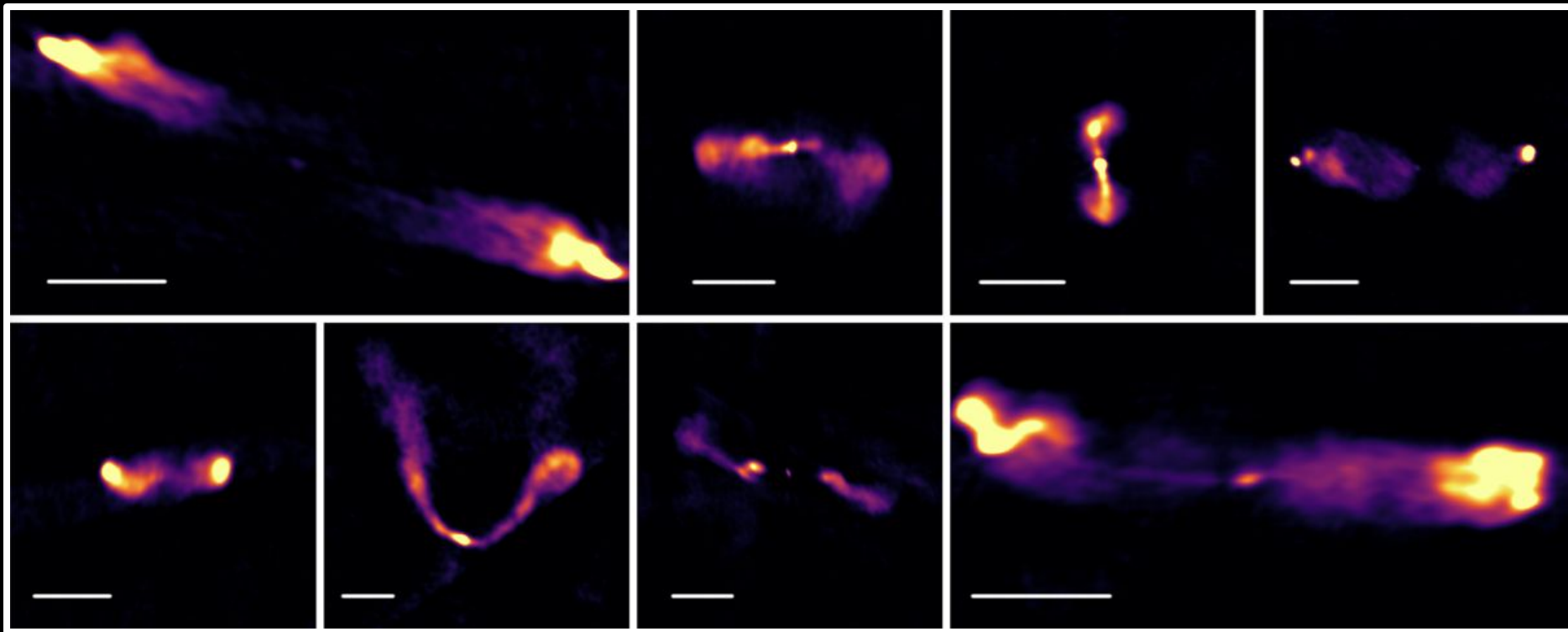
```
git push -u origin master
```

#Take a look and see what happened!

# Data Rates in Astronomy



- The Atacama Large Millimeter Array (ALMA) in Chile produced:
  - over 1300 Terabytes of total data in 2014.
  - over 2700 Terabytes of total data in 2016
- The Very Large Array in New Mexico has the capability of producing a million simultaneous frequency channels.
- VLA Sky Survey - over 300 GB ASDM for a 4 hour scheduling block
- VLASS will produce over 330 terabytes of imaging



# 3D Graphics Software



MAYA



3DS MAX



# HOUDINI

# 3D Graphics, Python, and Astronomy

I use a non-traditional package called Blender to render different forms of astronomical data - catalogs, data cubes, simulations, etc.







# What is Blender?

Blender is:

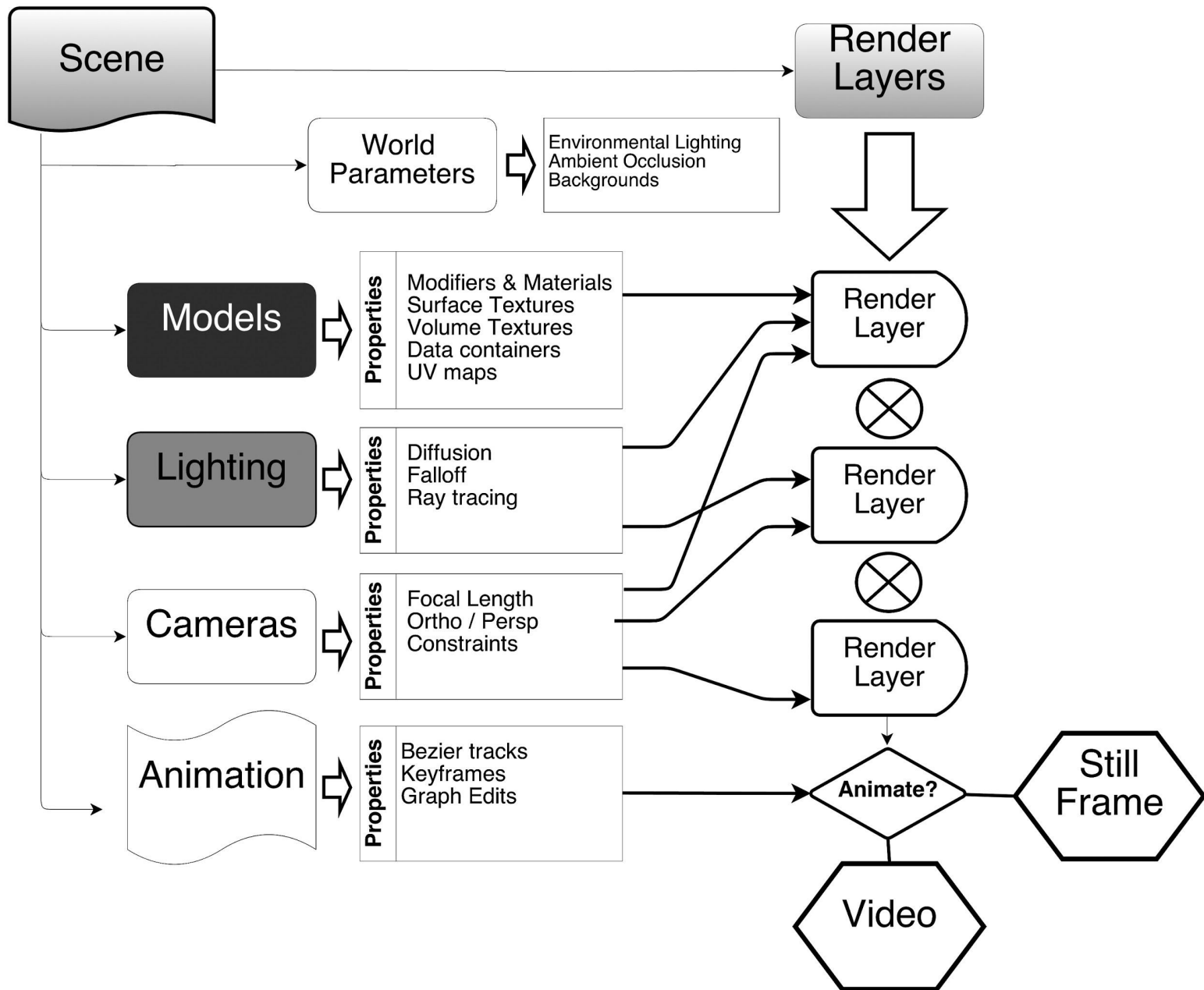
- 3D graphics software for modeling, animation, and visualization
- Open-source
- a real-time 3D viewer and GUI
- A Python scriptable interface for loading data

<http://www.blender.org>

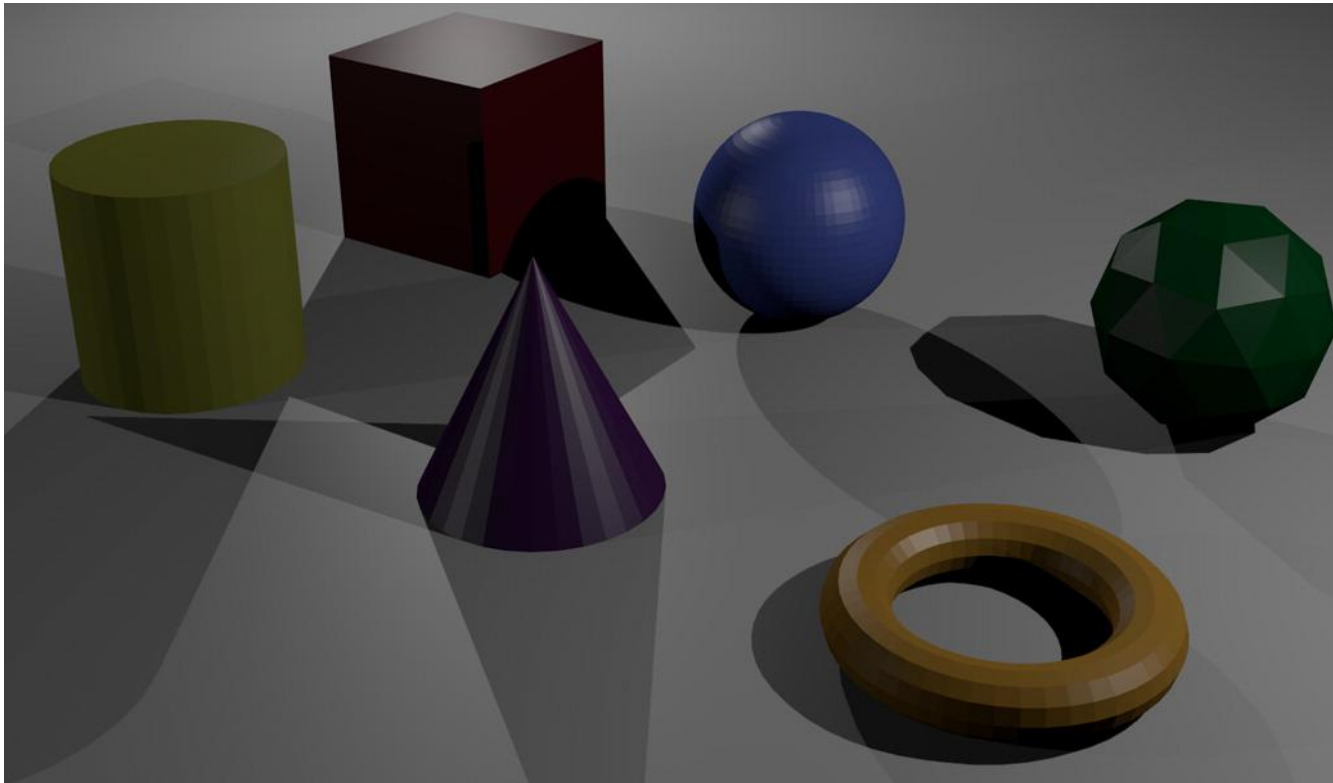
# Elements of 3D Graphics

We need to consider:

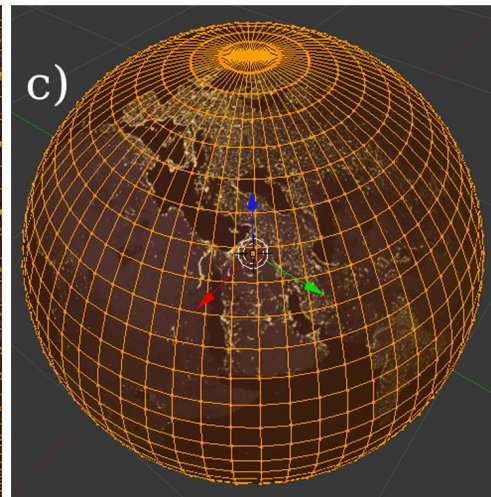
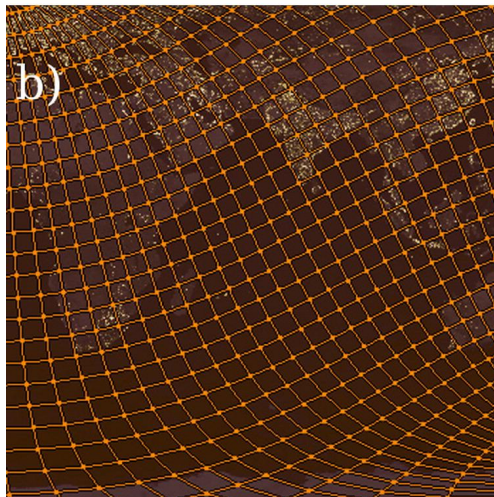
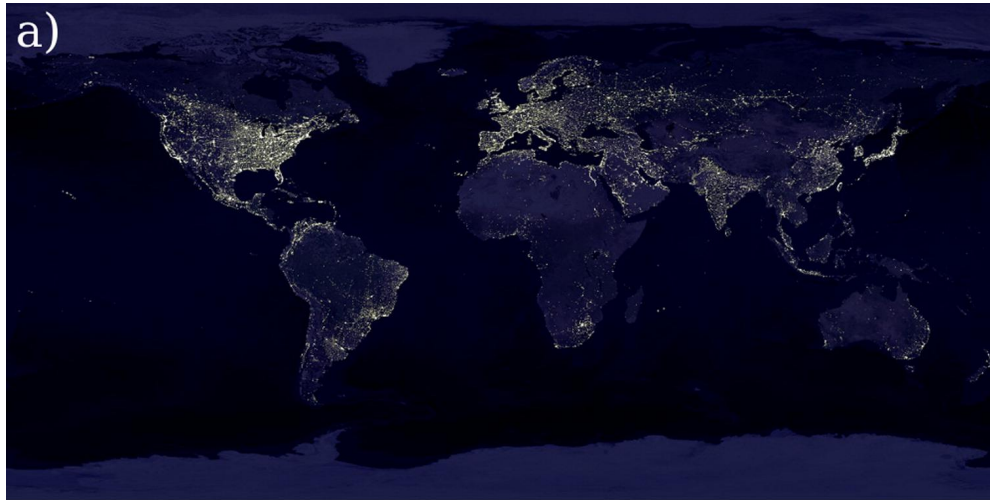
- Models - physical or data containers?
- Textures - 2D, 3D, and projections?
- Lighting - illumination of data - physical or artistic
- Animation - How will the model move and change?
- Camera control - lens selection, angle, image size, and movement and tracking
- Rendering - backend engine choice
- Compositing - layering final output



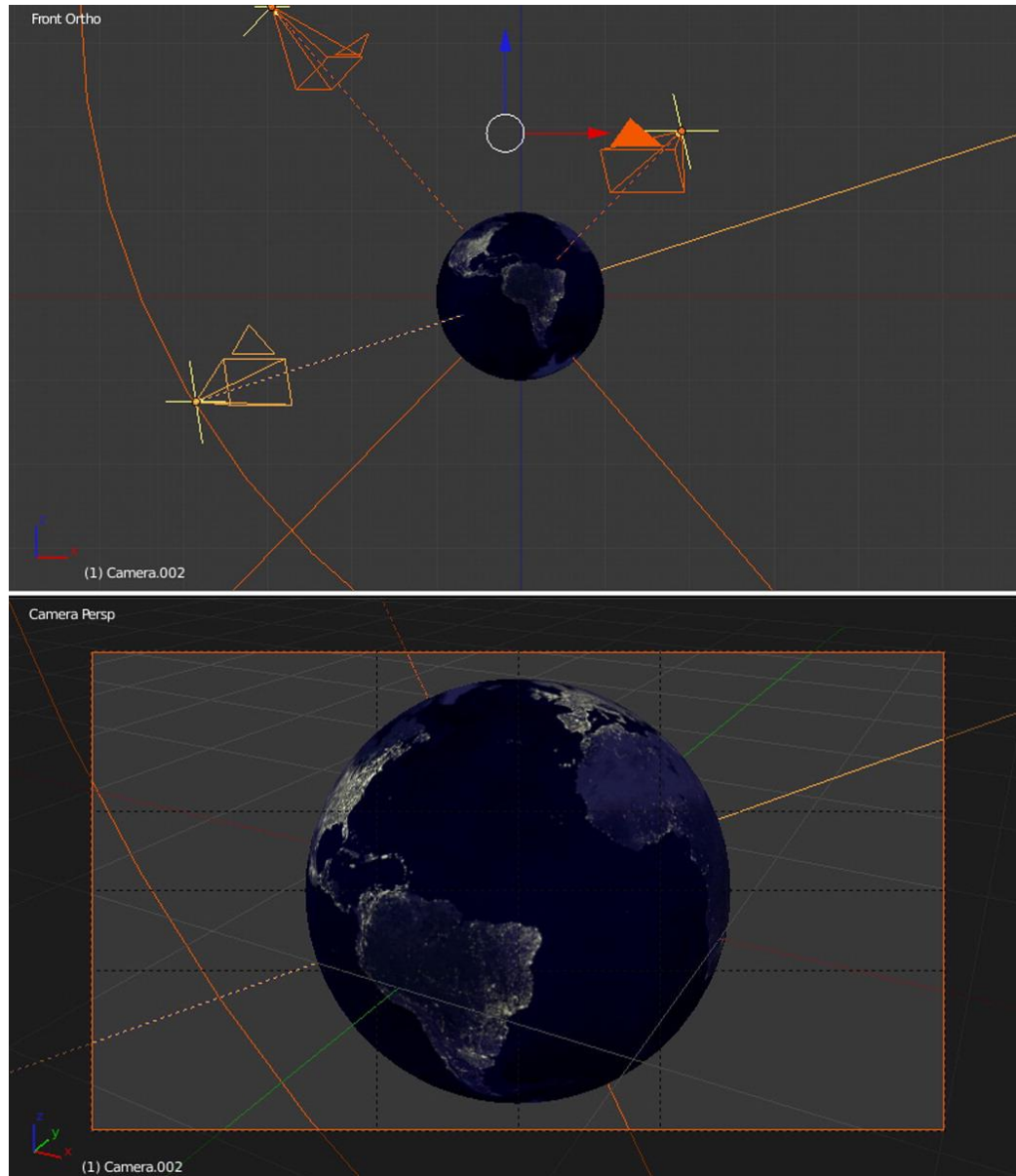
# Modeling - basic shapes and containers



# Texturing and Mapping



# Camera Control and Movement



Track View - Dope Sheet

Modes Controller Tracks Keys Time Options Display View Utilities

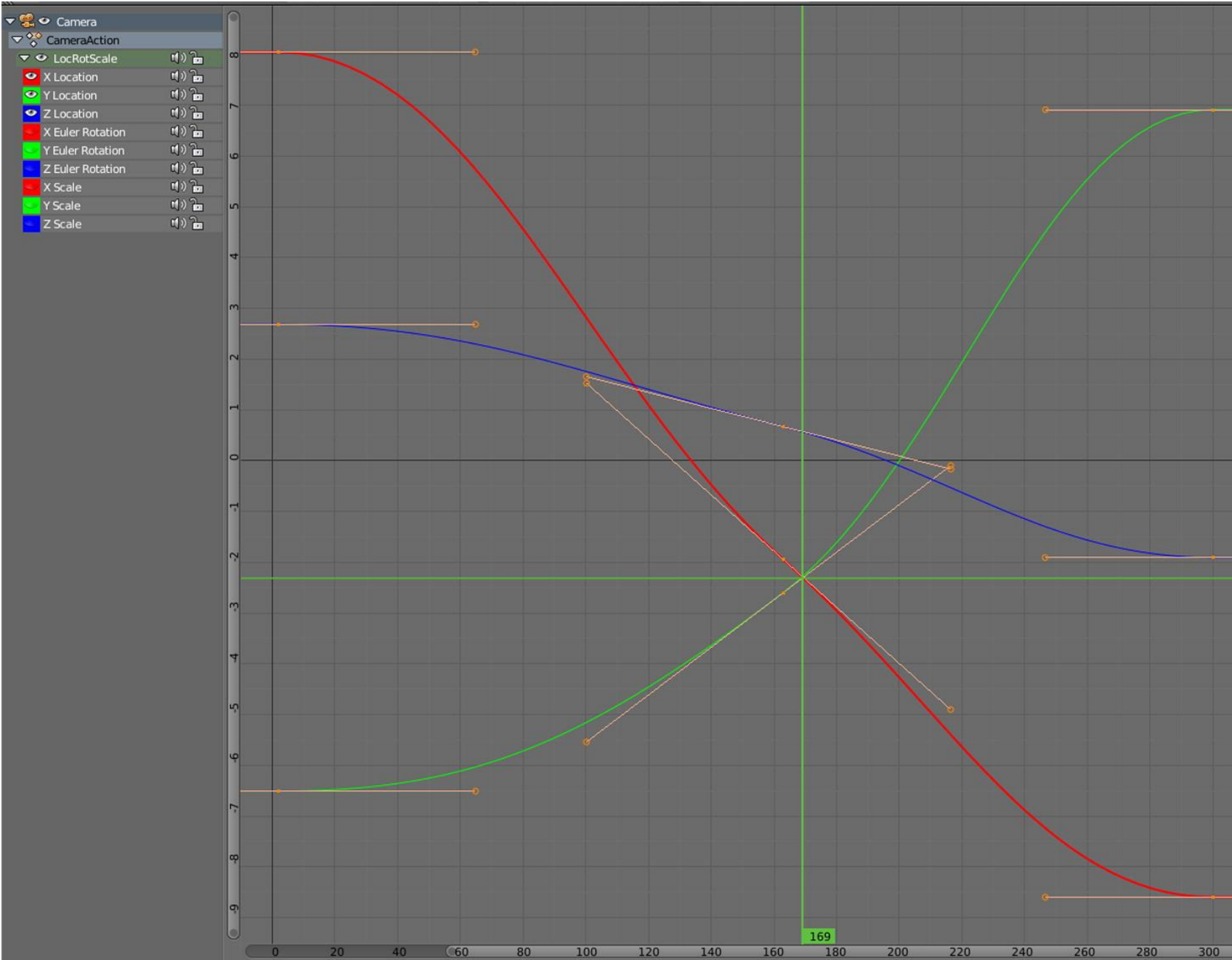
Box001

- Transform : Position/Rotation/Scale
  - Position : Position XYZ
    - X Position : Bezier Float
    - Y Position : Bezier Float
    - Z Position : Bezier Float
  - Rotation : Euler XYZ
    - X Rotation : Bezier Float
    - Y Rotation : Bezier Float
    - Z Rotation : Bezier Float
  - Scale : Bezier Scale
- Object (Box)

X: 100.0 Y: 100.0 Z: 100.0

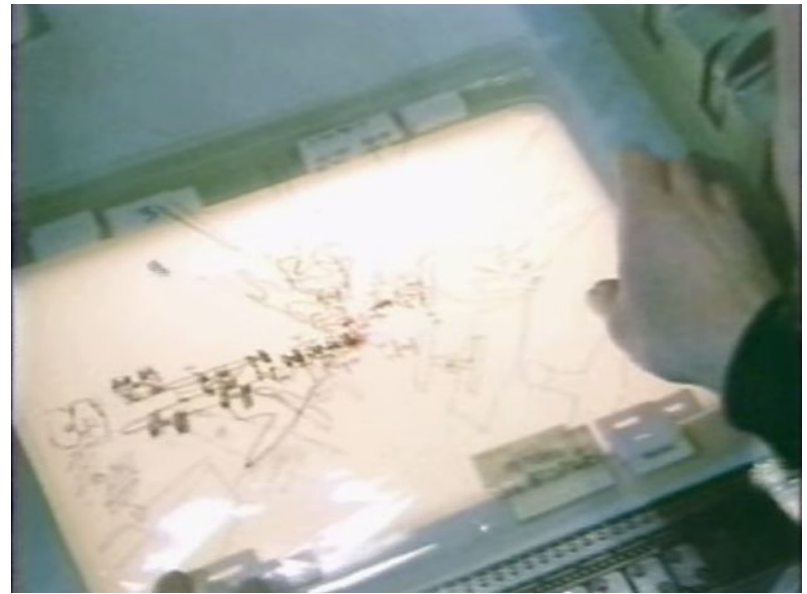
17 -26.295 4.2







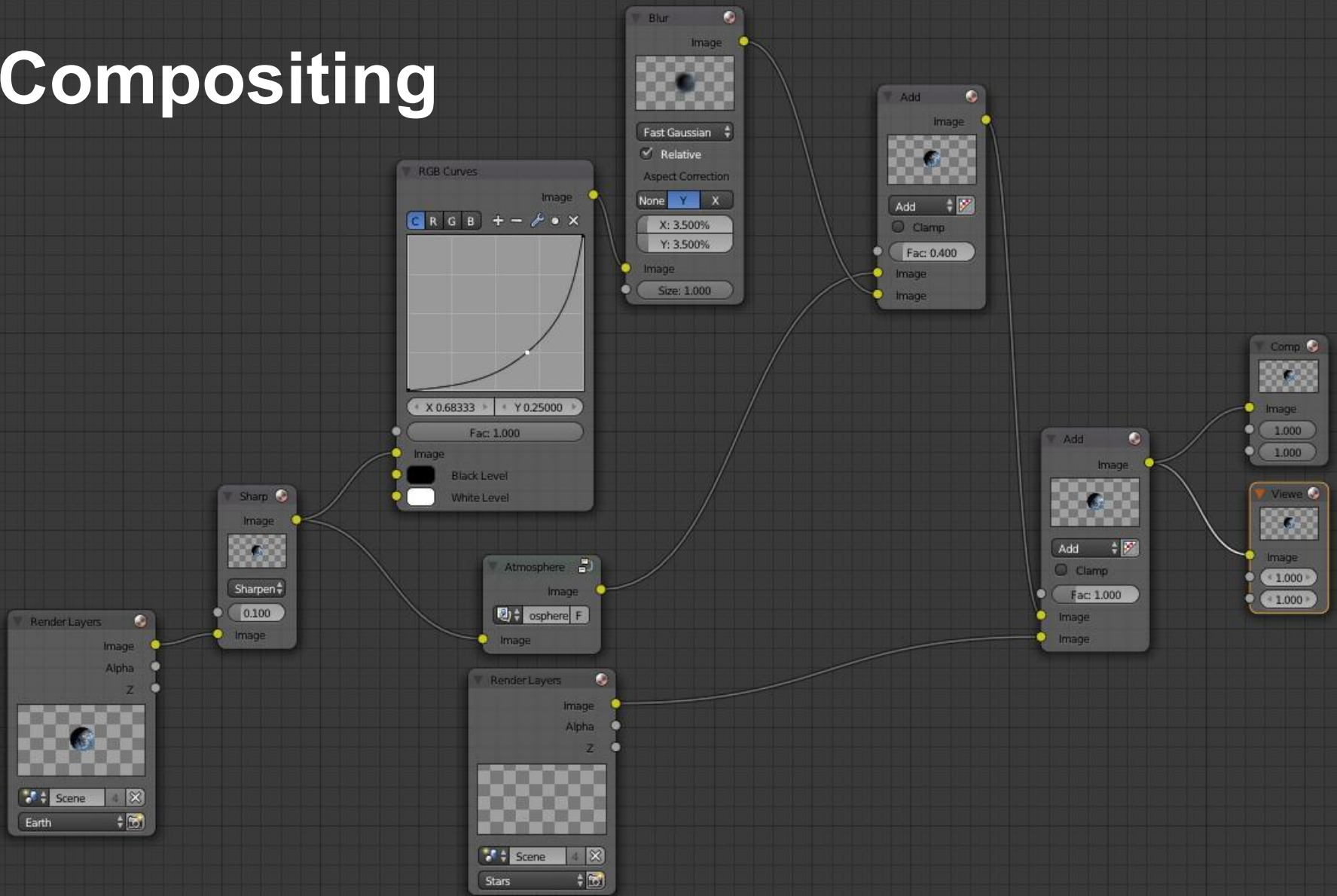
# Animation



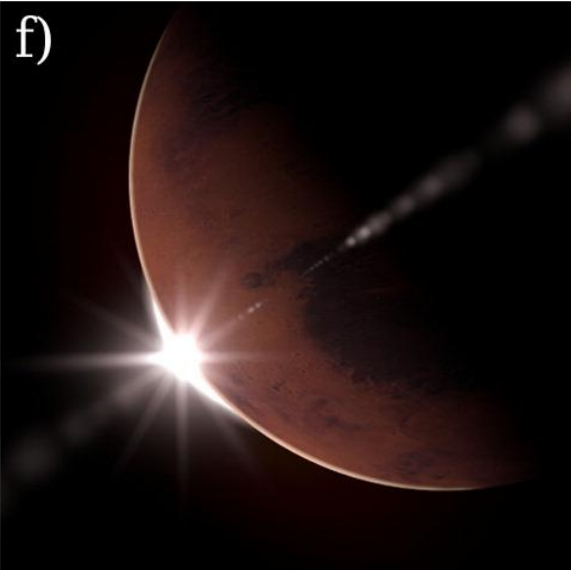
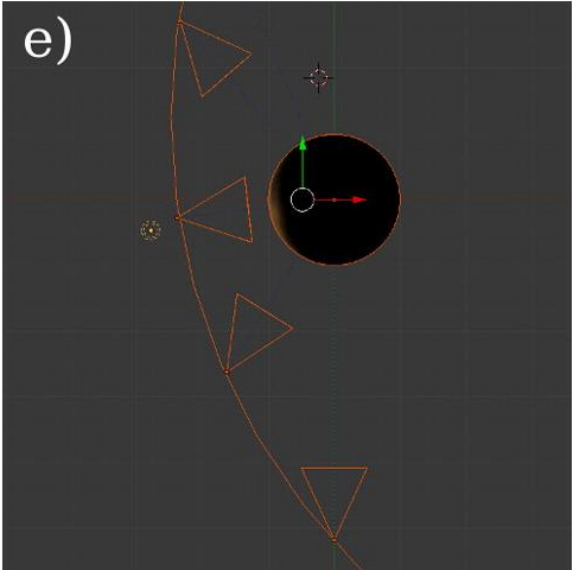
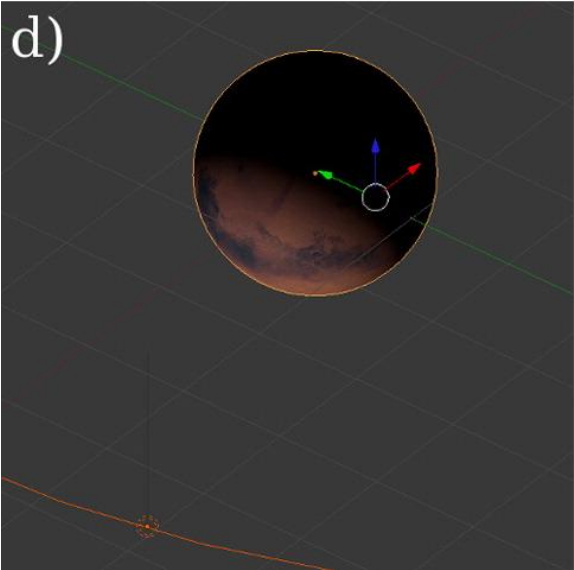
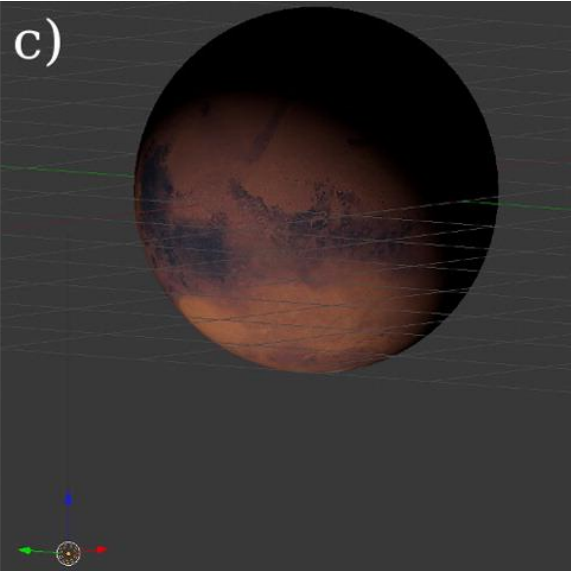
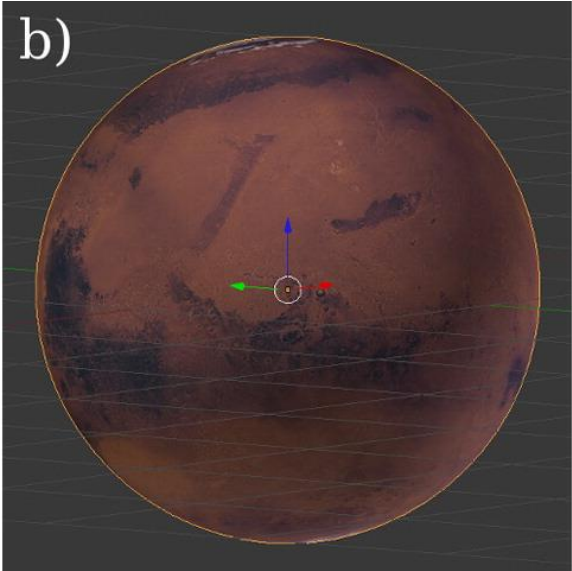
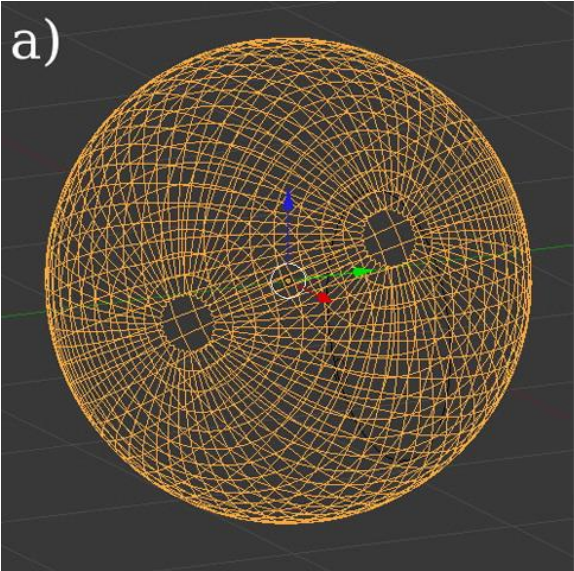
# Rendering Engine

- Blender (included)
- Cycles (included)
- Yafaray (open source ray tracing engine  
<http://www.yafaray.org/>)
- Luxrender ([http://www.luxrender.net/en\\_GB/index](http://www.luxrender.net/en_GB/index))
- Octane (<http://render.otoy.com/>)
- Renderman (<http://renderman.pixar.com/view/renderman>)

# Compositing



# Rendering and Compositing

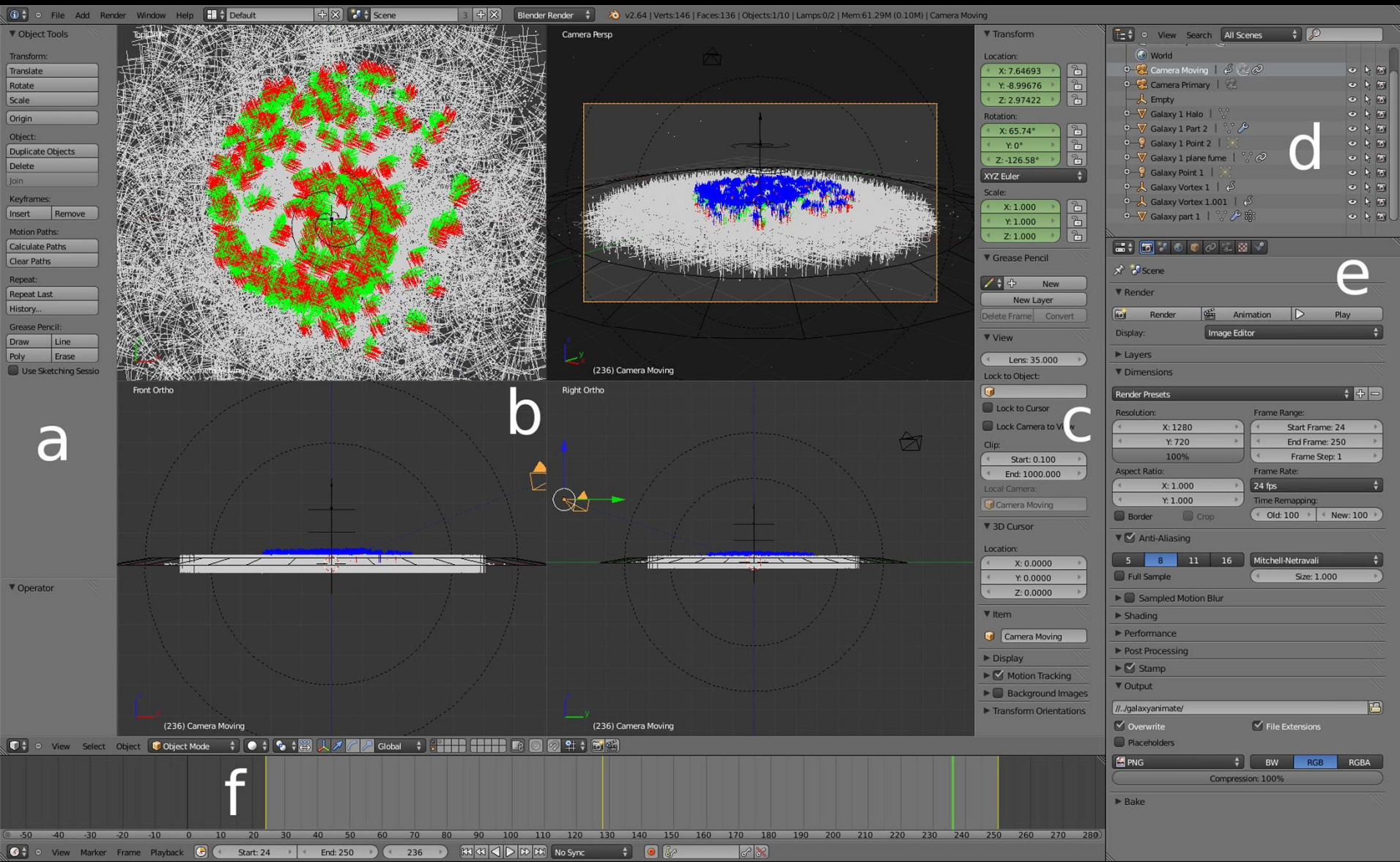




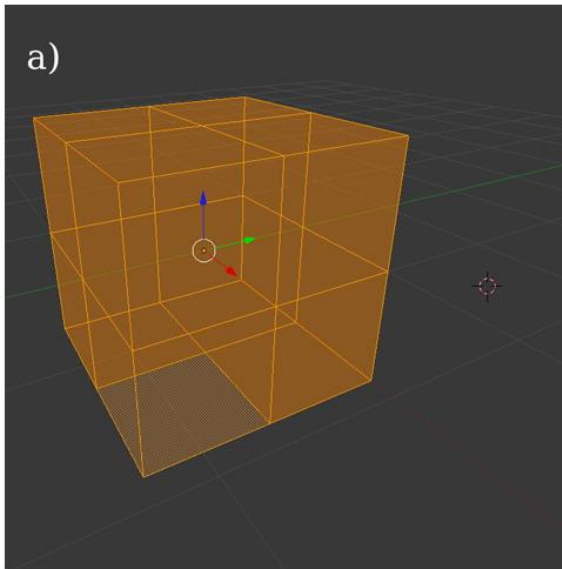


# Examples

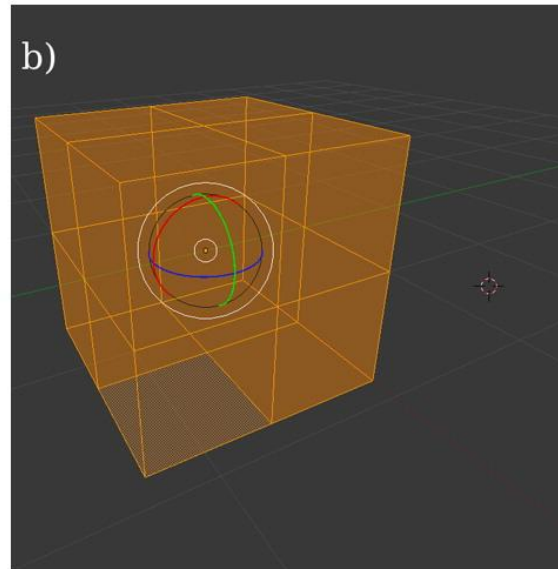
# **A Tour of the Blender Interface**



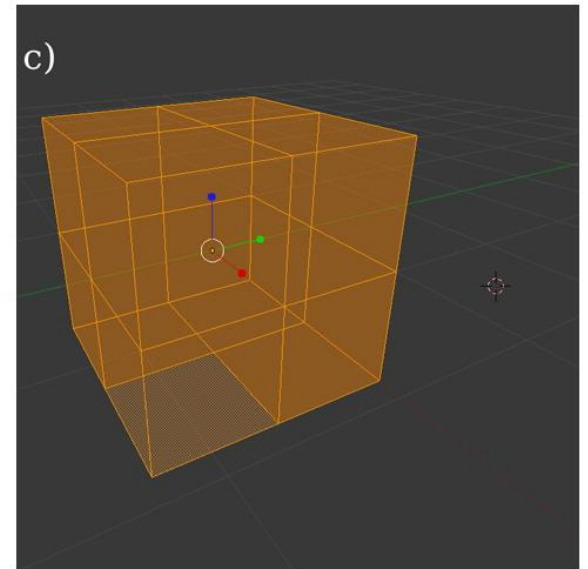
# Blender interface



**Translation**



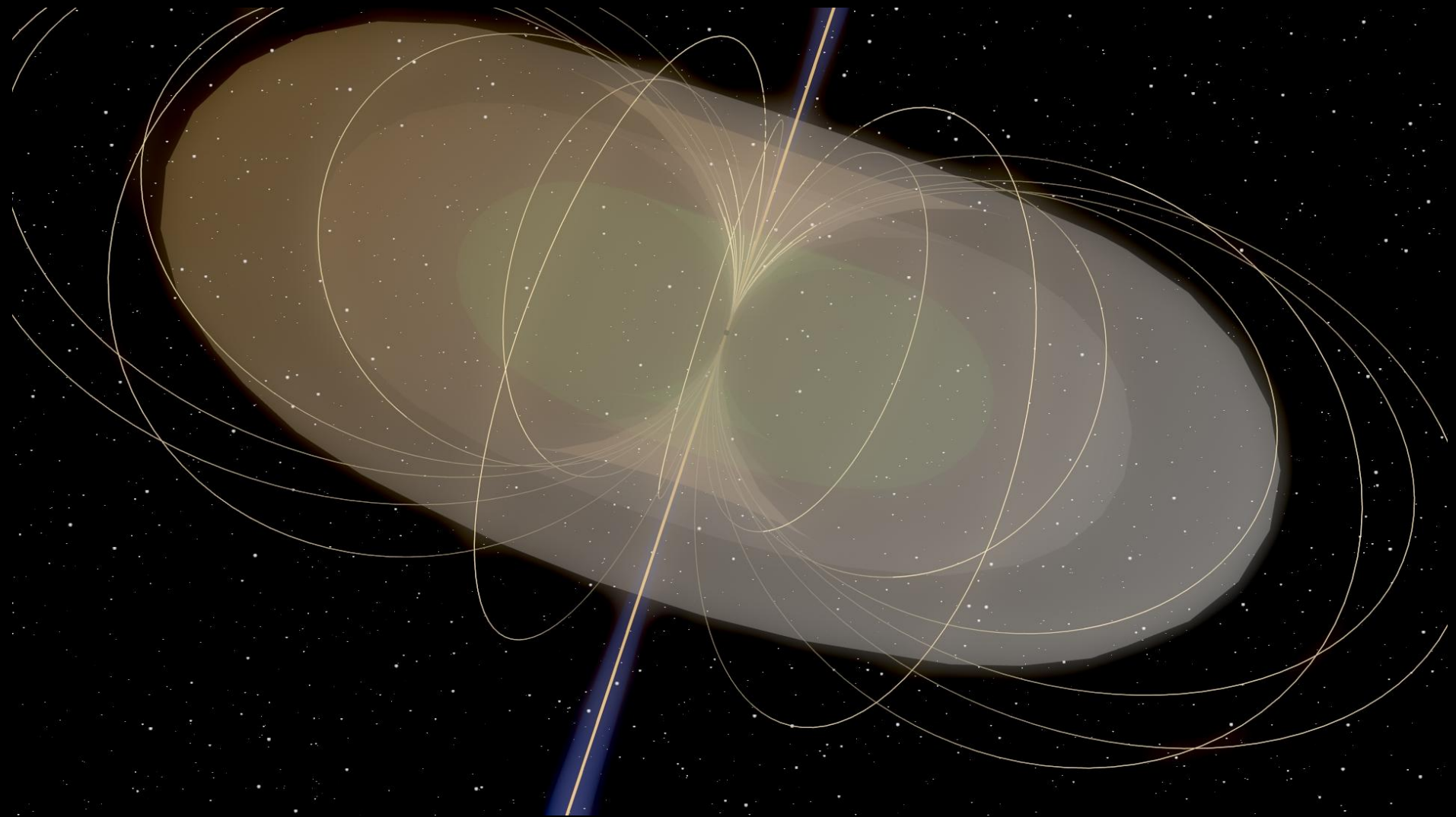
**Rotation**



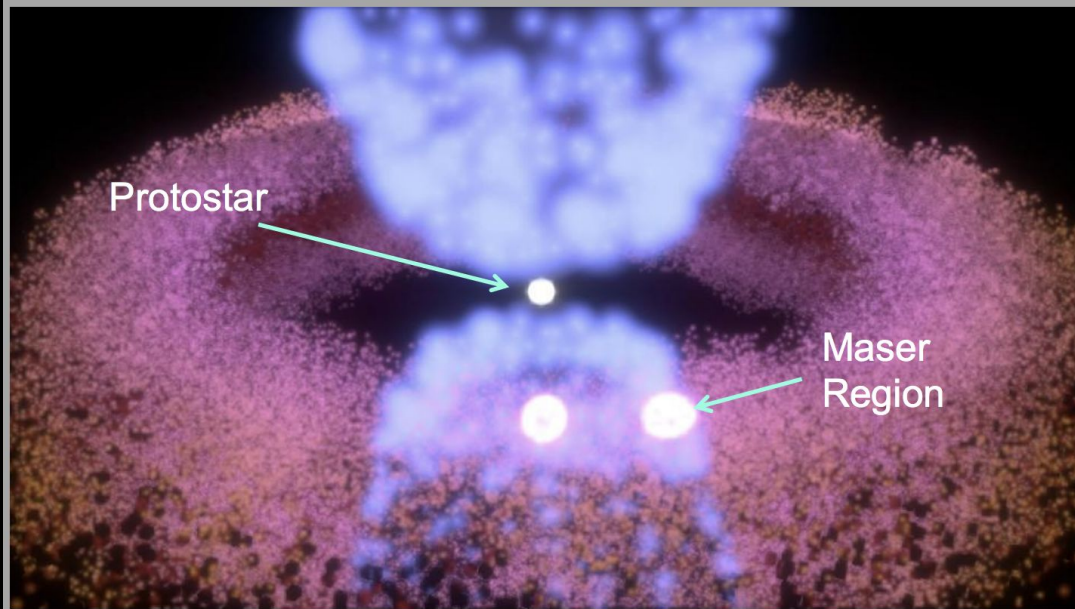
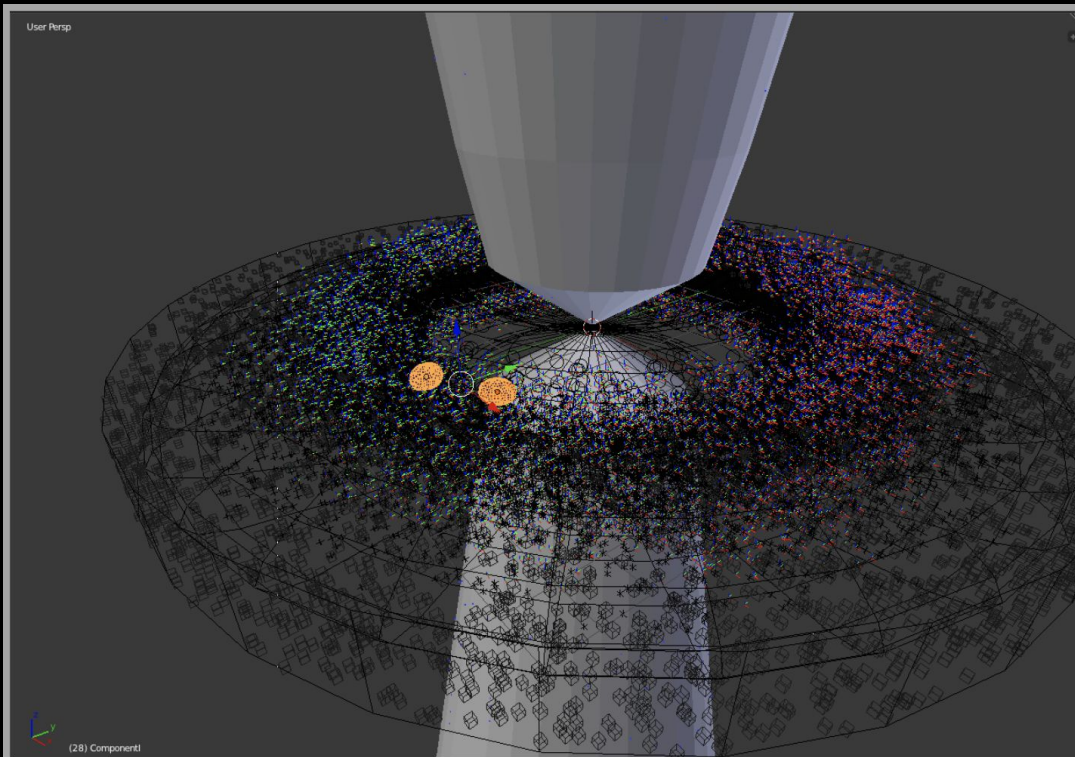
**Scaling**



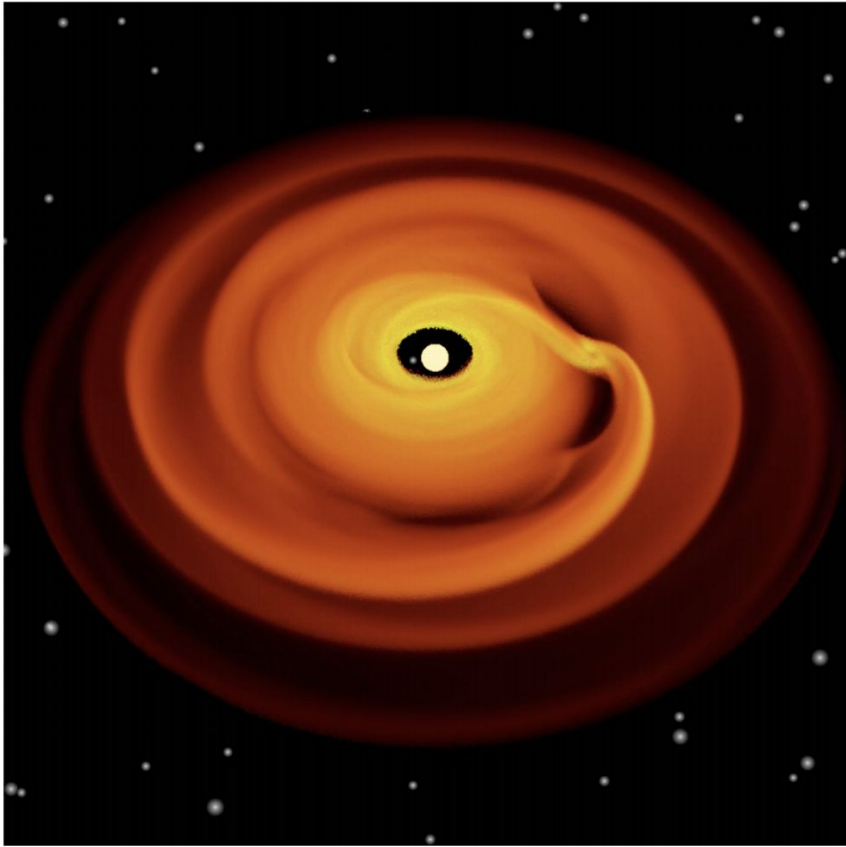






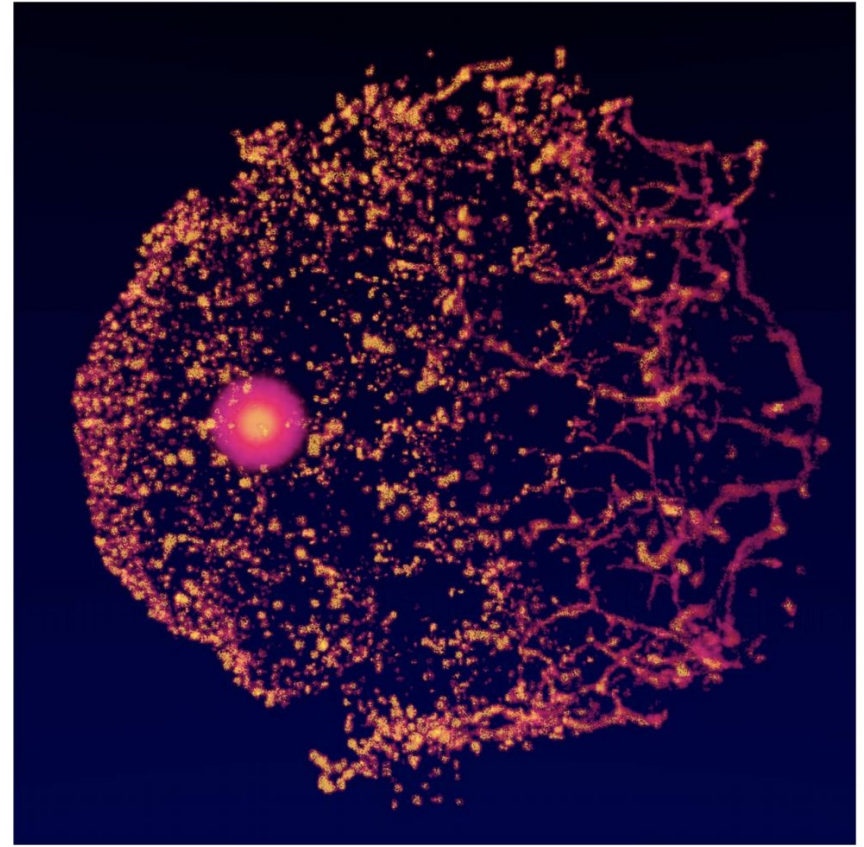


Araya et al.  
Western Illinois Univ.



**Figure 4.** Protoplanetary disk with a massive planet carving a gap (S. Perez et al. 2017, in preparation). Simulation data provided by S.Perez using FARGO3D. The output was converted from an spherical grid as described in the Section 3.1. The image was post-processed to brighten the colors, and the halo points were added to emulate surrounding stars.

(A color version of this figure is available in the online journal.)



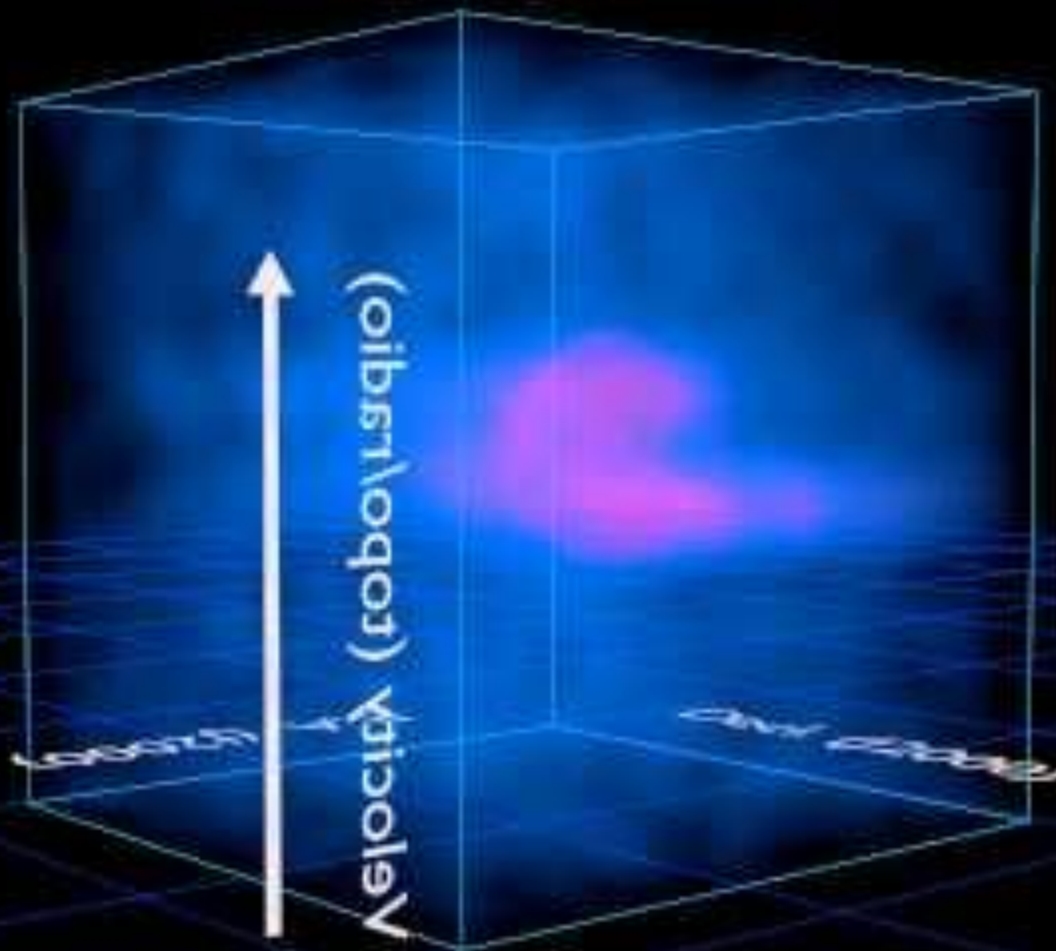
**Figure 5.** Stellar winds from a Wolf-Rayet star moving through the galactic center (Cuadra et al. 2008). Simulation data provided by J. Cuadra. The output was converted from an SPH simulation as described in Section 3.2. The image was post-processed to brighten the colors, and also the Blend Sky option of the World properties was used to add the background colors.

(A color version of this figure is available in the online journal.)

# Data Cubes

- Gridded data can come from telescopes or simulations
- Radio telescopes produce grids that cover...
  - Two sky coordinates (RA and Decl.)
  - Frequency (Z - the doppler shifted velocity)
- These cubes can show the dynamics of galaxies, planetary disks, and large scale structure formation of clusters

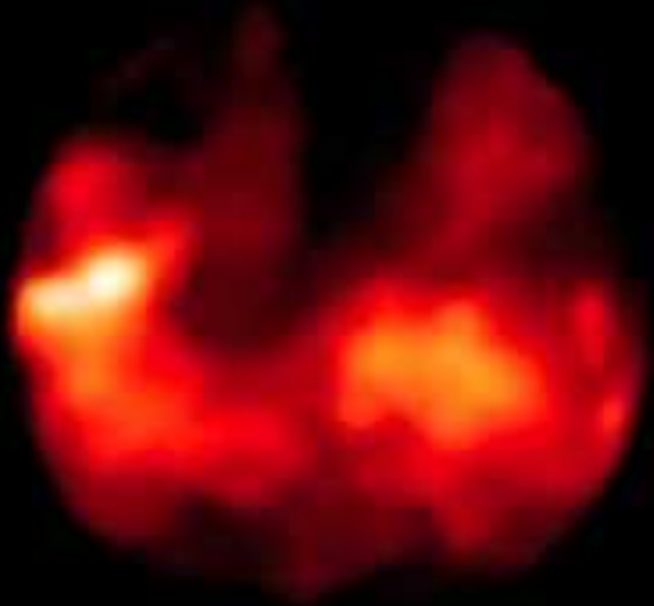
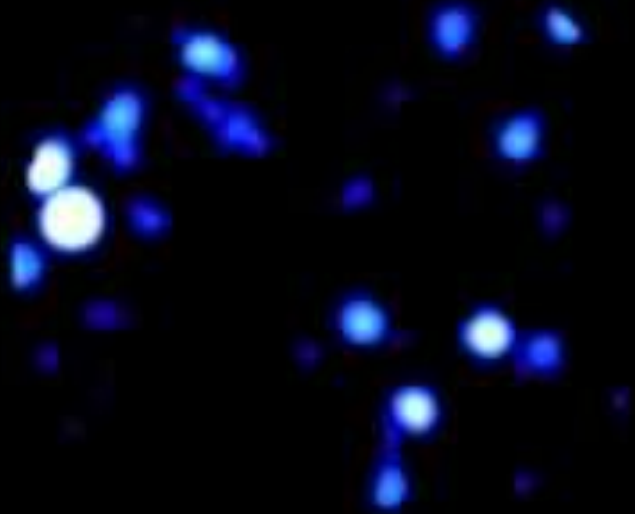




<https://www.youtube.com/watch?v=RDUVZ9MIW2I>

# Data Cubes

- Density maps of the nearby Universe can be created on regularly spaced grids.
- The results of these surveys allow to study not only the density of galaxies in 3D, but also the effects of gravity in the same regions of space...

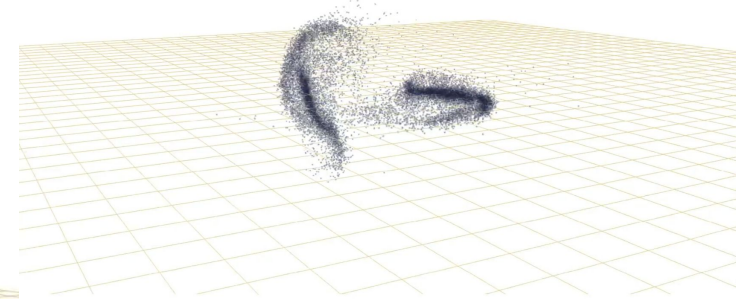
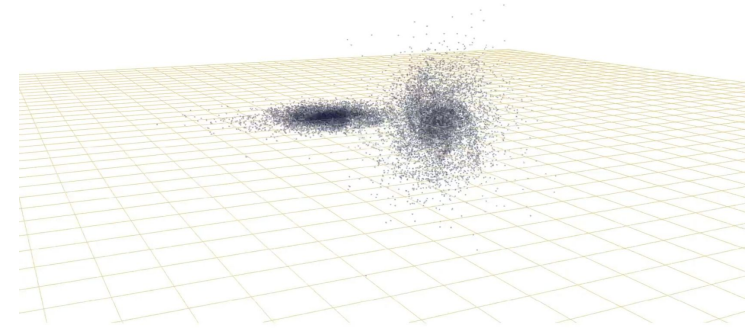
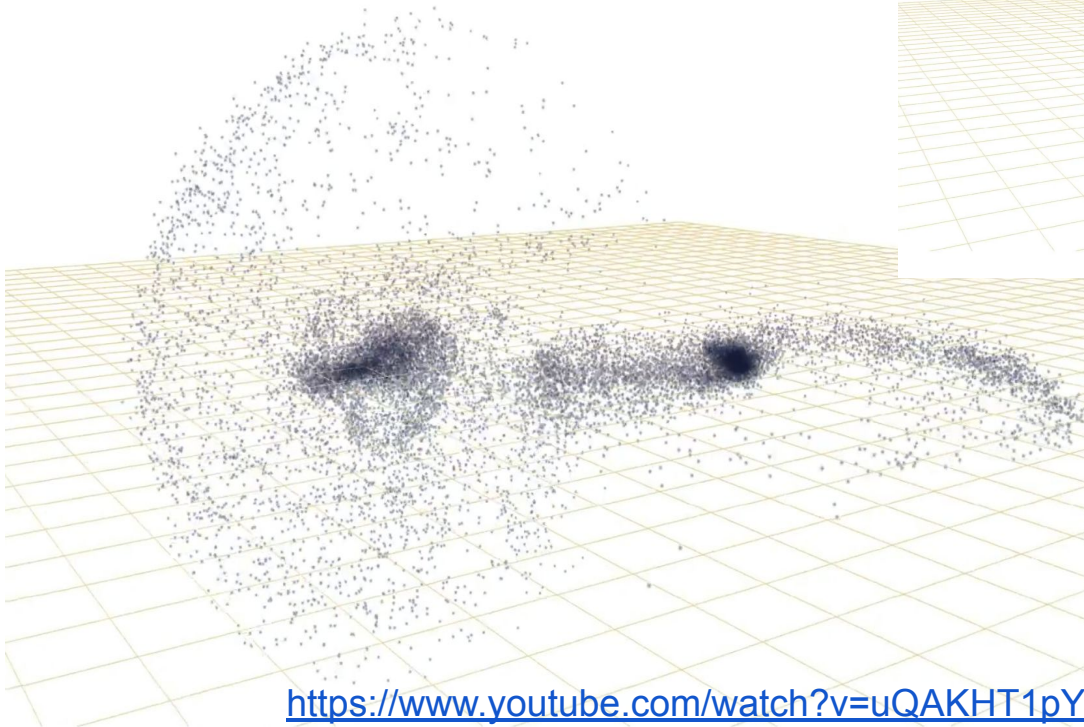


[https://www.youtube.com/watch?v=3cuNT8\\_YEF0](https://www.youtube.com/watch?v=3cuNT8_YEF0)



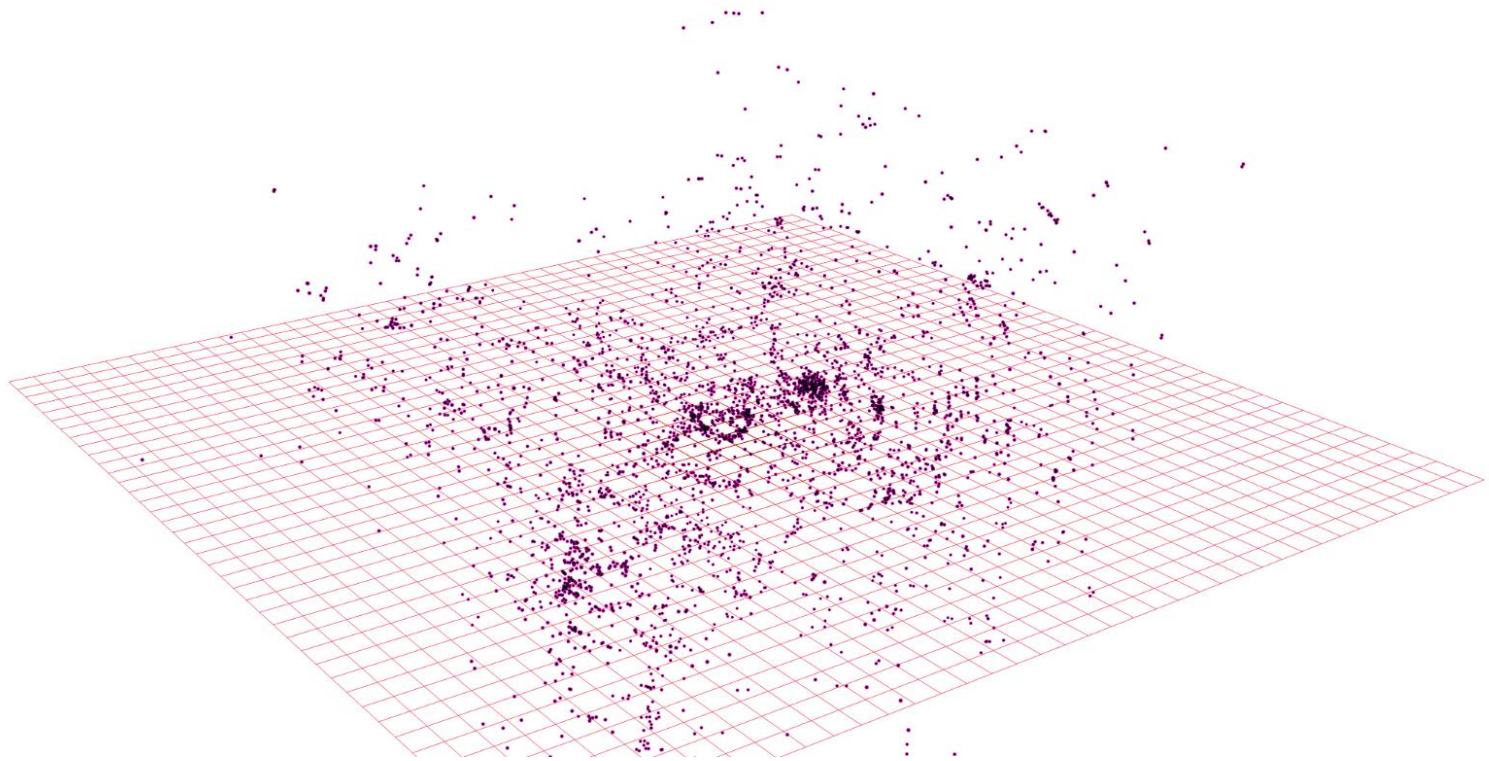
# N-body Simulations

- Data generated from GADGET-2 (Galaxies and Dark Matter Interacting 2) N-body/SPH code
  - <http://www.mpa-garching.mpg.de/gadget/>
- 30,000 particles, 1100 snapshots run for 2 billion years
- Blender Python interface used to bring XYZ position data into the vertices of Blender objects
- Objects are “textured” with Halos.
- Each grid square is approximately 33,000 light years



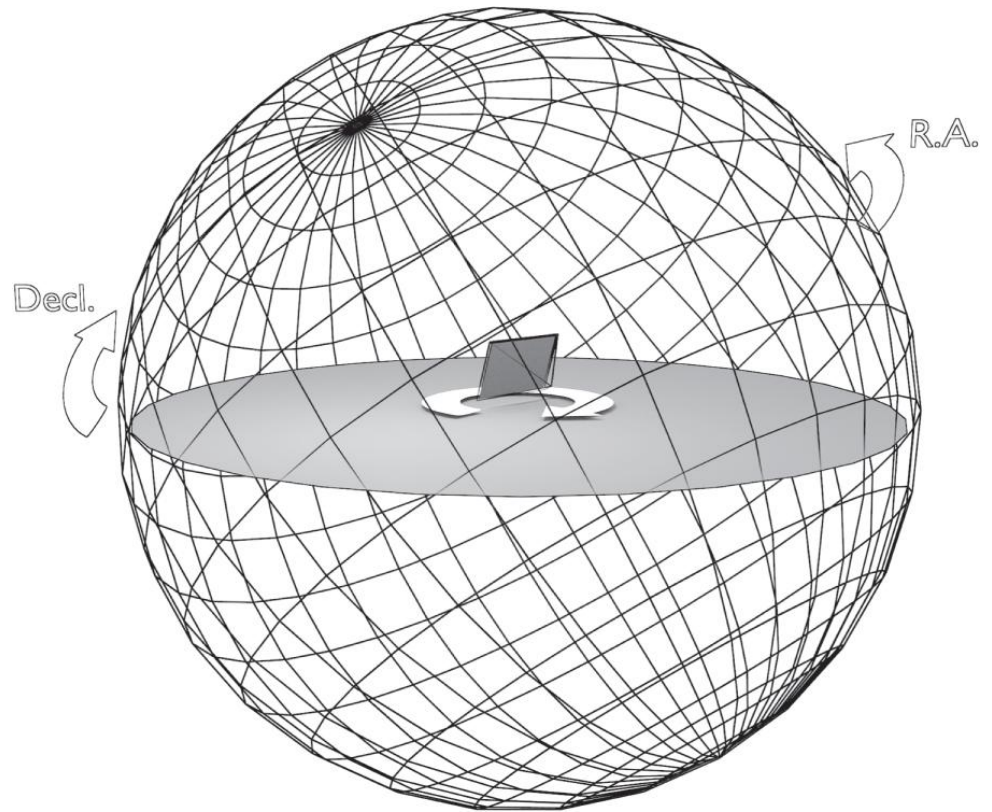
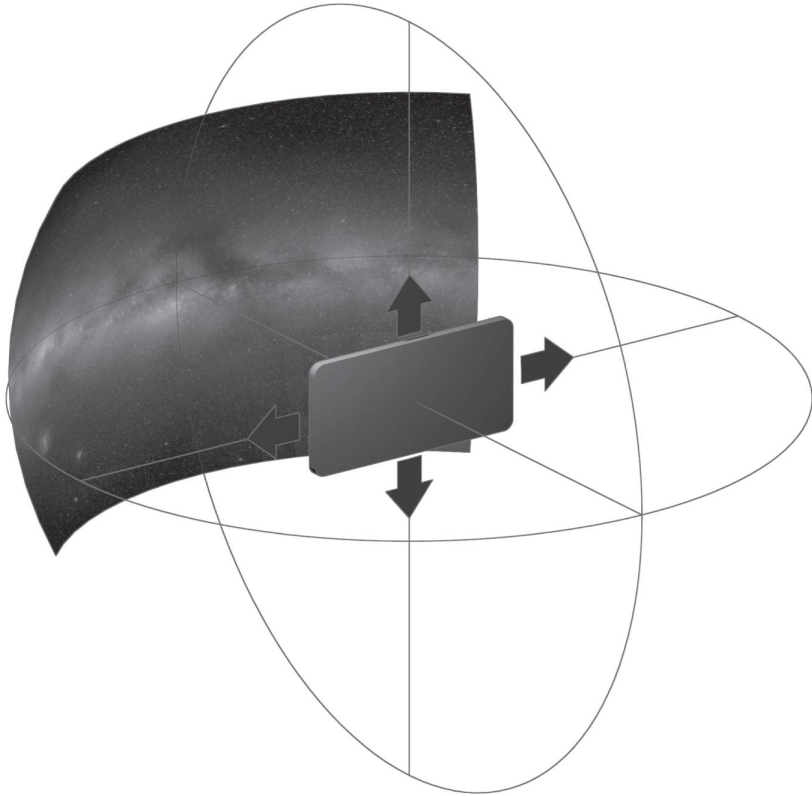
<https://www.youtube.com/watch?v=uQAKHT1pY9s>

# Galaxy Catalogs

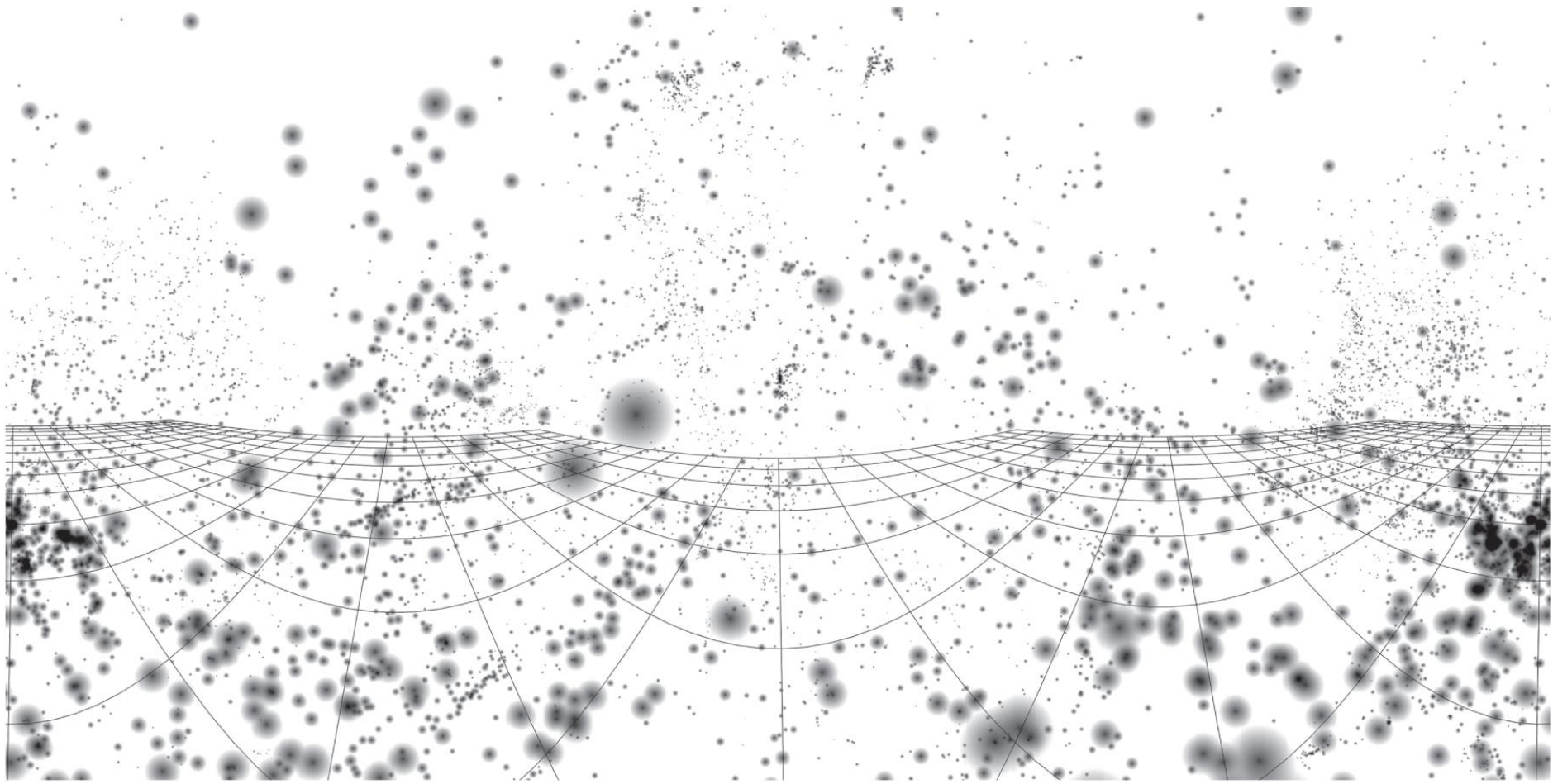


[https://www.youtube.com/watch?v=eO\\_OKyfDAyM](https://www.youtube.com/watch?v=eO_OKyfDAyM)

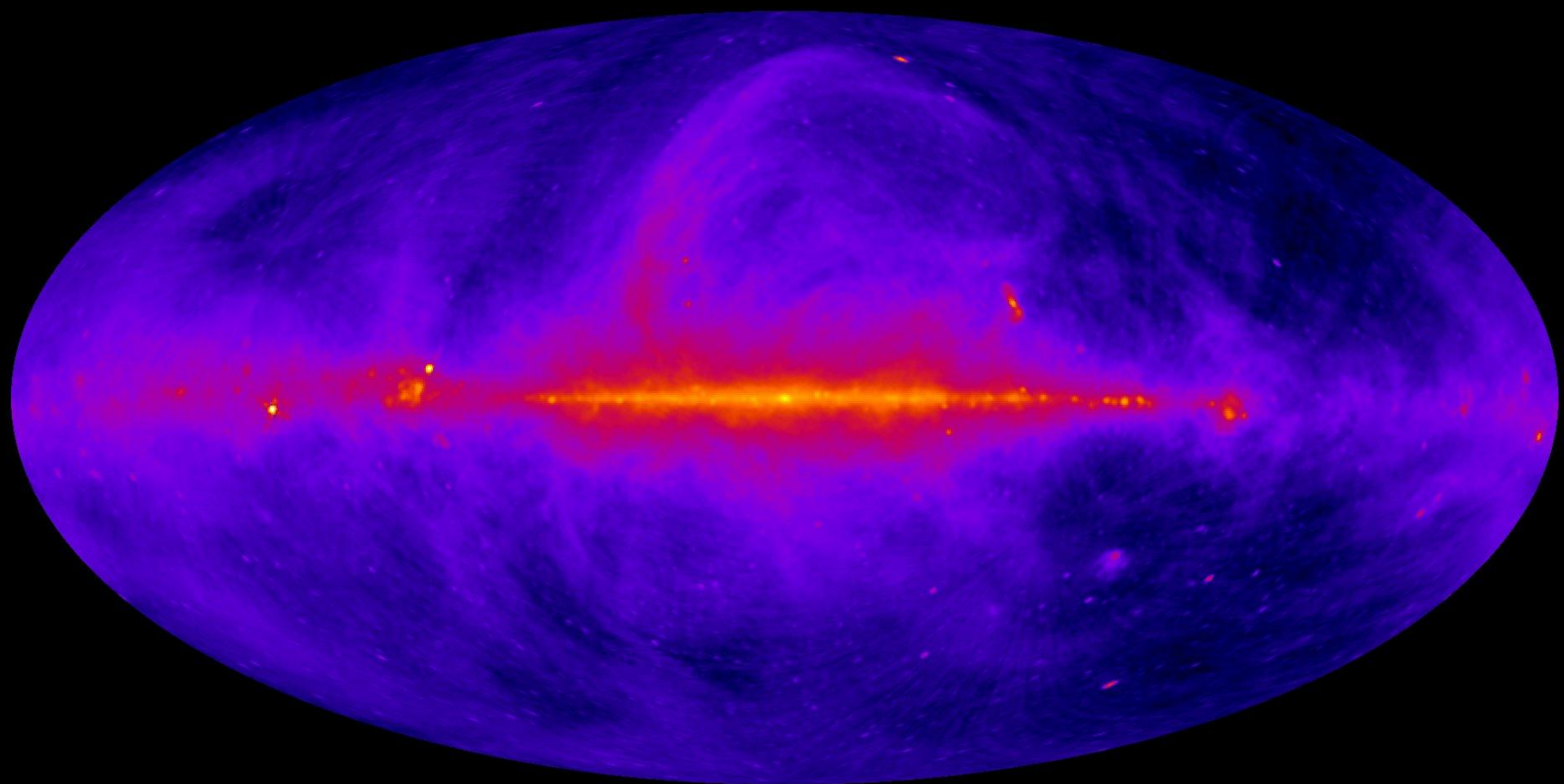
# 360 Panoramas







Tully et al. Cosmic Flows data



<https://www.youtube.com/watch?v=YWwA49Mm1nw>

408 MHz

## Techniques and Methods for Astrophysical Data Visualization

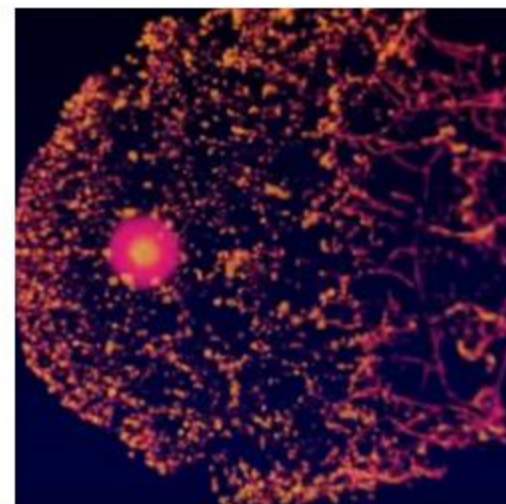
**Brian R. Kent**, *National Radio Astronomy Observatory, Charlottesville, VA, USA*

Astrophysics continues to be a leader in the data sciences, with innovative methods being developed to handle new analysis challenges. The higher rates of data acquisition in both observational and theoretical astrophysics demand innovative solutions in scientific visualization. The [Publications of the Astronomical Society of the Pacific \(PASP\)](#) has published a special focus issue titled **Techniques and Methods for Astrophysical Data Visualization**. Refereed submissions for this issue cover a wide variety of visualization topics, including new software packages, visualization techniques, software from other industries, and new science results. These methods and techniques can serve as a complement to data analysis software, stand on their own for data exploration, or inspire with impressive visuals for science, technology, education, mathematics (STEM) and public outreach. A number of articles from our special focus issue feature videos and tutorials as well as interactive 3D content.



"Visualization allows astronomers to break down and understand large data and explore multi-dimensional phase spaces. We encourage scientists to explore the tools and techniques presented in this issue and apply them to their own data and research."

- **Brian R. Kent, Guest Editor, PASP**



**Figure.** Stellar winds from a Wolf-Rayet star moving through the galactic center (Garate 2017). Simulation data provided by Cuadra et al. (2008).

# Interesting in learning more?



Book and tutorials  
available at:

<http://www.cv.nrao.edu/~bkent/blender/>

<https://www.youtube.com/VisualizeAstronomy>



# Resources and Opportunities

## 3D Graphics and Python

<http://www.cv.nrao.edu/~bkent/blender/>

## Blender

<http://www.blender.org/>

## General Visualization

<http://visual.ly/>

<http://www.visualizing.org/>

## SIGGRAPH and Pycon

<http://www.siggraph.org>

<https://www.youtube.com/channel/UCsX05-2sVSH7Nx3zuk3NYuQ> - Pycon 2018

