



**SEAMLESS  
ASTRONOMY**

Linking scientific data, publications, and communities

# Astronomy Dataverse:

*enabling astronomer data publishing*

<http://theastrodata.org>



# SEAMLESS ASTRONOMY

Linking scientific data, publications, and communities



## Harvard-Smithsonian Center for Astrophysics



HARVARD  
LIBRARY



The  
**Dataverse  
Network**<sup>TM</sup>  
Project



An Open-Source Application for  
Publishing, Citing and Discovering Research Data



# LECTURES

*De Potentia Restitutiva,*

OR OF

# SPRING

Explaining the Power of Springing Bodies.

# in the past data were hidden...

About two years since I printed this Theory in an Anagram at the end of my Book of the Descriptions of Helioscopes, viz. *ceiinossttuv*, id est, *Ut tensio sic vis*; That is, The Power of any Spring is in the same proportion with the Tension thereof: That is, if one power stretch or bend it one space, two will bend it two, and three will bend it three, and so forward. Now as the Theory is very short, so the way of trying it is very easie.

1660 Robert Hooke “pre” published anagram:

- “*ceiinossttuv*”
- “*ut tensio, sic vis*”
- *as the tension, so the force*

# in the present data live in papers

The screenshot shows a web browser window with the URL <http://iopscience.iop.org/ezp-prod1.hul.harvard.edu/0067-0049/166/1/249/fulltext>. The page is from 'THE ASTROPHYSICAL JOURNAL SUPPLEMENT SERIES' and features the title 'Hubble Space Telescope Proper Motions and Stellar Dynamics in the Core of the Globular Cluster 47 Tucanae<sup>1</sup>' by Dean E. McLaughlin, Jay Anderson, Georges Meylan, Karl Gebhardt, Carlton Pryor, Dante Minniti, and Sterl Phinney. The abstract text is visible, starting with 'We have used HST imaging of the central regions of the globular cluster 47 Tucanae (≡NGC 104), taken with the W instruments between 1995 and 2002, to derive proper motions and U- and V-band magnitudes for 14,366 stars with core radii of the cluster center. This represents the largest set of member velocities collected for any globular cluster involved range in brightness from just fainter than the horizontal branch of the cluster to more than 2.5 mag below the turnoff. In the course of obtaining these kinematic data, we also use a recent set of ACS images to define a list of astrophysical positions (and F475W magnitudes) for nearly 130,000 stars in a larger, ≈3' × 3' central area. We describe procedures in some detail and provide the full position, photometric, and velocity data in the form of downloadable FITS files. We have used the star counts to obtain a new estimate for the position of the cluster center and to define the density sequence turnoff and giant branch stars in to essentially zero radius, thus constraining the global spatial structure of the cluster. A single-mass, isotropic King model fit to it is then used as a rough point of reference against which to compare the characteristics of our proper-motion data. We search in particular for any evidence of very fast-moving stars, in sign numbers that expected for the extreme tails of the velocity distribution in a sample of our size. We find that likely few no more than about 0.3%, of stars with measured proper motions have total speeds above the nominal central escape velocity. At lower speeds, the proper-motion velocity distribution very closely matches that of a regular King model (nearly Gaussian given the high stellar density) at all observed radii. Considerations of only the velocity dispersion the number of results: (1) Blue stragglers in the core of 47 Tuc have a velocity dispersion  $\sigma_{\mu}$  smaller than that of the cluster factor of  $\sqrt{2}$ , consistent with the former being on average twice as massive as normal, main-sequence turnoff stars. (2) The distribution in the inner 5 core radii of the cluster is essentially isotropic, and the detailed dependence of  $\sigma_{\mu}$  on  $R$  for suggests that heavy remnants contribute only a fraction of a percent to the total cluster mass. Both of these results are earlier, more realistic multimass and anisotropic models of 47 Tuc. (3) Using a sample of 419 line-of-sight velocities bright giants within  $R \leq 105''$ , we obtain a kinematic distance to the cluster,  $D = 4.0 \pm 0.35$  kpc, formally some 10% recent estimates based on standard color-magnitude diagram fitting, and more consistent with the value implied by dwarf cooling sequence. And (4) by fitting simple models of isotropic, single-mass stellar clusters with central point observed  $\sigma_{\mu}(R)$  profile, we infer a  $1\sigma$  upper limit of  $M_{\bullet} \lesssim 1000-1500 M_{\odot}$  for any intermediate-mass black hole in the best-fit hole mass ranges from zero, if only the kinematics of stars near the main-sequence turnoff mass are modeled if fainter, less massive stars are also used. We can neither confirm nor refute the hypothesis that 47 Tuc might lie on the  $M_{\bullet}-\sigma$  relation observed for galaxy bulges.

Subject headings: [astrometry](#); [globular clusters: individual \(NGC 104\)](#); [stellar dynamics](#)

Online material: [FITS file](#), [machine-readable table](#), [tar file](#)

<sup>1</sup> Based on observations made with the NASA/ESA Hubble Space Telescope, obtained at the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy (AURA), Inc., under NASA contract NA48-002-26400.

Online Material  
FITS file

▶ [Master frame reference image](#)

▶ [Machine-readable table](#)

▶ [Table 4](#)

▶ [Tar file](#)

▶ [Proper-motion catalog](#)

References: McLaughlin et al. 2006; <http://adsabs.harvard.edu/abs/2006ApJS..166..249M>

# Tables, Tables in tar file

Search Criteria

Keywords

- J/ApJS/166/249

Tables

- J/ApJS/166/249
- ..table3
- ..table4
- ..table5

Table Name	Description
<input type="checkbox"/> J/ApJS/166/249/table3	(c) Stars used for astrometric calibration
<input type="checkbox"/> J/ApJS/166/249/table4	(c) 129733 stars in the Master Frame
<input type="checkbox"/> J/ApJS/166/249/table5	(c) Proper motions and displacements

Reset All

# Code in tar file

## FITS Files

File: 477tocMaster.f

Object Value: 124

FK5

RA: 00 24 09.803

Dec: -72 04 47.29

Physical

X: 2619.483

Y: 3107.644

Image

X: 2619.483

Y: 3107.644

Frame 1

Zoom: 0.666

Angle: 0.000

```
1 #
2 # MACRO pmdat REQUIRES ONE COMMAND-LINE ARGUMENT ...
3 #
4 # $1 = ID number of a star in file '...datfile'
5 #   in either of two formats: M11111 = an exact ID label in the file
6 #   : 11111 = integer part [>0] of a known ID
7 # -- OR --
8 #
9 # $1=0 or $1=[any string not starting with 'M']
10 #   will choose a star at random from the file '...datfile'
11 #
12 #
13 # MACRO pmdat WILL OPTIONALLY TAKE A SECOND ARGUMENT...
14 #
15 # $2 = PRINT [optional]
16 #
17 #   if second argument exists and is PRINT
18 #   then star data are printed to file with extension '.DATA'
19 #
20 #   if second argument is anything else, or does not exist at all,
21 #   then star data are echoed to screen instead
22 #
23 #
24 #
25 # Usage... at the sm prompt, type
```

**References:** McLaughlin et al. 2006; <http://adsabs.harvard.edu/abs/2006ApJS..166..249M>



VizieR

# Tables, Tables in tar file

new CDS Portal : search by position for available data in CDS services (Simbad, Aladin)

Stellar dynamics and proper motion

J/ApJS/166/249

<input type="checkbox"/>	J/ApJS/166/249/table3	(c) Stars used for astrometric calibration
<input type="checkbox"/>	J/ApJS/166/249/table4	(c) 129733 stars in the Master Frame
<input type="checkbox"/>	J/ApJS/166/249/table5	(c) Proper motions and displacements

Reset All



# FITS Files

# Code in tar file

```

1 #
2 # MACRO pmdot REQUIRES ONE COMMAND-LINE ARGUMENT ...
3 #
4 # $1 = ID number of a star in file "...datafile..."
5 #   in either of two formats: M:llll = an exact ID label in the file
6 #   : llll = integer part (>=0) of a known ID
7 # -- 0 ---
8 #
9 # $! = 0 if $1=[any string not starting with 'M']
10 #   will choose a star at random from the file "...datafile..."
11 #
12 #
13 # MACRO pmdot [ID] [OPTIONALLY TAKE A SECOND ARGUMENT...]
14 #
15 # $2 = PRINT [optional]
16 #
17 #   if second argument exists and is PRINT
18 #   then star data are printed to file with extension ".DATA"
19 #
20 #   if second argument is anything else, or does not exist at all,
21 #   then star data are echoed to screen instead
22 #
23 #
24 #
25 # Usage... at the sm prompt, type

```

References: McLaughlin et al. 2006; <http://adsabs.harvard.edu/abs/2006ApJS..166..249M>



VizieR



# Tables, Tables in tar file

Table Name	Description
<input type="checkbox"/> J/ApJS/166/249/table3	(c) Stars used for astrometric calibration
<input type="checkbox"/> J/ApJS/166/249/table4	(c) 129733 stars in the Master Frame
<input type="checkbox"/> J/ApJS/166/249/table5	(c) Proper motions and displacements

**FITS Files**

**Code in tar file**

```

1 #
2 #
3 #
4 #
5 #
6 #
7 #
8 #
9 # $! = 0
10 # $! = [any string not starting with 'W']
11 # will choose a star at random from the file "...datafile"
12 #
13 # MACRO pndot [OPTIONALLY TAKE A SECOND ARGUMENT...]
14 #
15 # $2 = PRINT [optional]
16 #
17 # If second argument exists and is PRINT
18 # then star data are printed to file with extension ".DATA"
19 #
20 # If second argument is anything else, or does not exist at all,
21 # then star data are echoed to screen instead
22 #
23 #
24 #
25 # Usage... at the sm prompt, type

```

**Unlinked Data is LOST data!**

References: McLaughlin et al. 2006; <http://adsabs.harvard.edu/abs/2006ApJS..166..249M>





```

1 #
2 # MICRO print: REQUIRES ONE COMMAND-LINE ARGUMENT ...
3 #
4 # $1 = ID number of a star in file "_datafile"
5 # in either of two formats: WU1111 = an exact ID label in the file
6 # "11111" = integer part [4] of a known ID
7 # -- OR --
8 #
9 # $1=# or $1=(only string not starting with "W")
10 # will choose a star at random from the file "_datafile"
11 #
12 #
13 # MICRO print: WILL OPTIONALLY TAKE A SECOND ARGUMENT ...
14 #
15 # $2 = PRINT [optional]
16 #
17 # If second argument exists and is PRINT
18 # then star data are printed to file with extension ".data"
19 #
20 # If second argument is anything else, or does not exist at all,
21 # then star data are echoed to screen instead
22 #
23 #
24 # Usage ... at the prompt, type
25 #

```

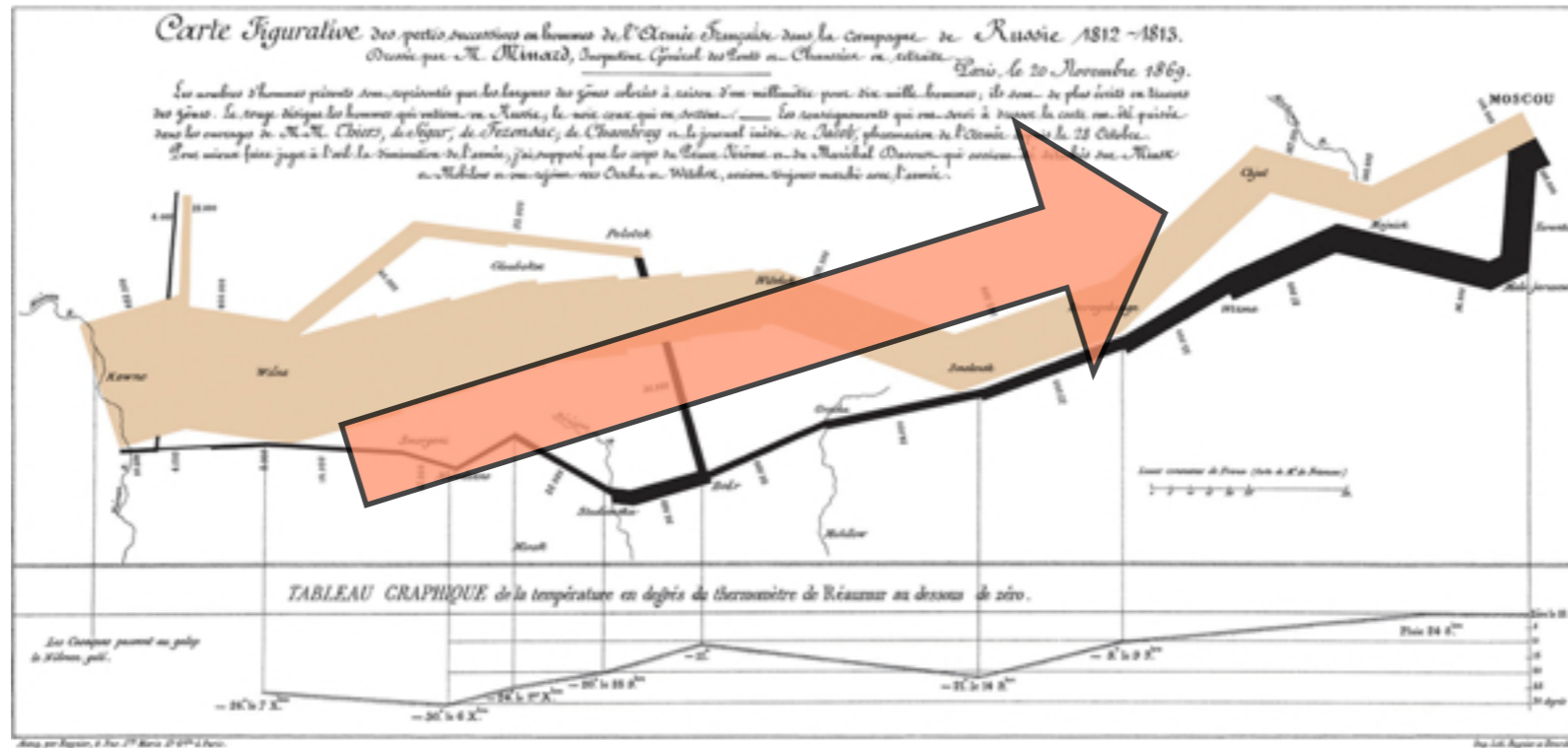
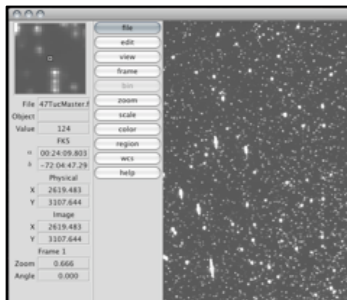


Table 2

WFC3 and ACS Observations of 47 Tucanae

Data Set	Program ID	N <sub>ex</sub>	Filter	Date
MEYLANG...	7842	15	F336W	1995 Oct 25 = 1995.82
MEYLANG...	6482	16	F300W	1997 Nov 3 = 1997.84
GILLIET...	6261	28	F336W	1999 Jul 5 = 1999.51
MEYLANG...	2863	16	F300W	1999 Oct 28 = 1999.82
GILLIET...	5206	11	F336W	2001 Jul 15 = 2001.53
WFC-MEUR...	9623	20	F437W	2002 Apr 5 = 2002.26
HEC-MEUR...	9623	40	F437W	2002 Apr 5 = 2002.26
HEC-BORG...	9623	10	F437W	2002 Apr 13 = 2002.28
WFC-KING...	7443	6	F437W	2002 Jul 7 = 2002.52
HEC-KING...	7443	20	F437W	2002 Jul 24 = 2002.56

3.3.3 Astrometric Calibration

We now have a position for the cluster center in the reference frame, which is based on the distance corrected and rotated frame of the first image of WFC3. In order to transform the cluster frame positions into absolute right ascension and declination, we used the image header information from several WFC3 images (GILLIET, MEYLANG, and GILLIET) to obtain absolute positions for seven stars—five stars at the center and two stars in the outskirts. These four images were taken at different parallaxes and orientations, so they should all use different parallax rates and give independent estimates of the absolute coordinates.

## Losses from Data to Literature

- Raw data:
  - ➡ might already be in a telescope archive
  - ➡ linkage partially fixed by post-pub curation
- Theoretical data;
- Analysis codes and logs;
- Processed data:
  - ➡ *Reduced data; mosaics;*

**References:** Charles Minard (1781-1870) (see upload log) [Public domain], via Wikimedia Commons



# in the future data live...

- Refined data sets are published by scientists in long lived repositories;
- Scientist's data linked in ADS & are “searchable”
- Scientist's data is **reused & cited**, giving credit for that work.



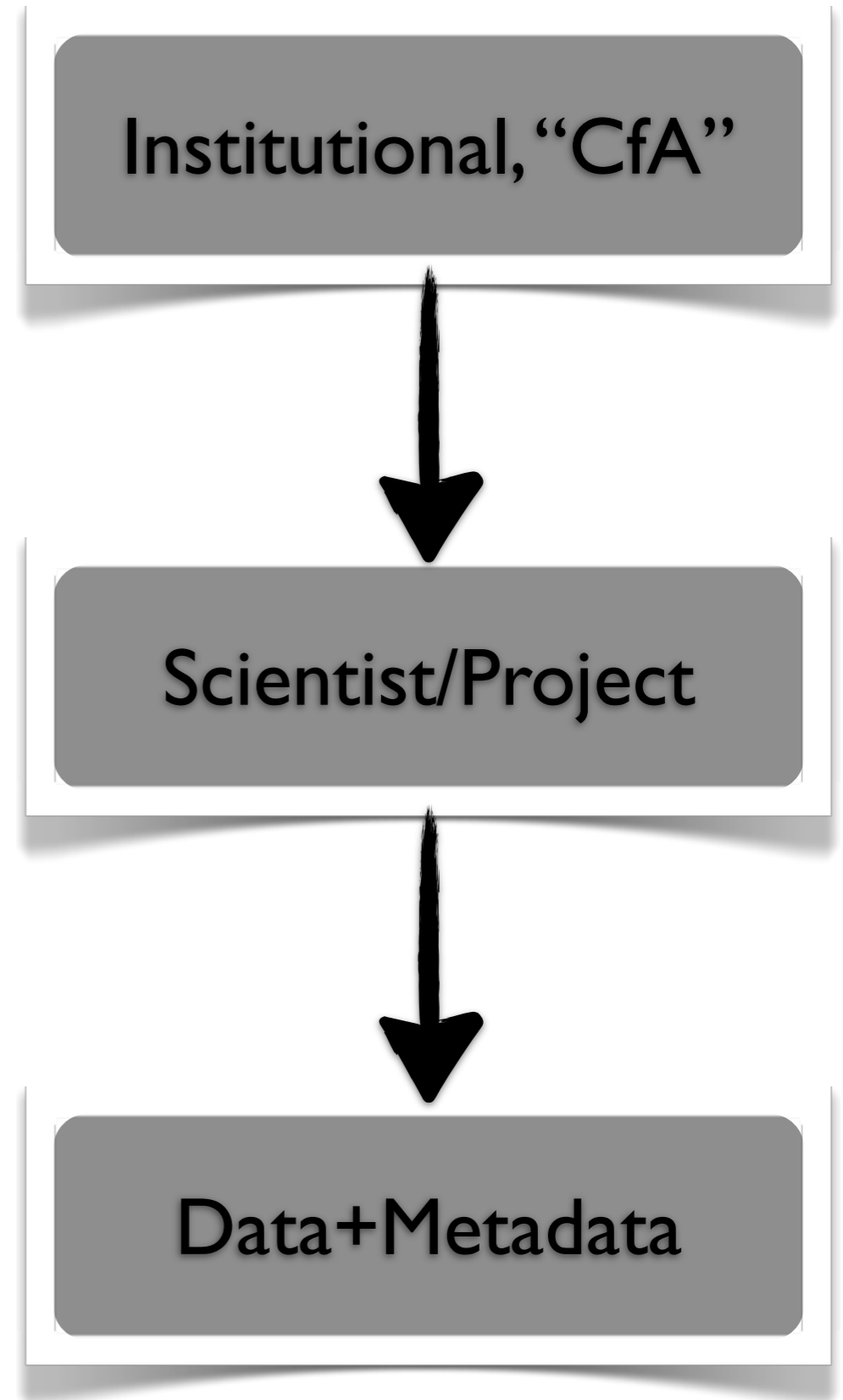
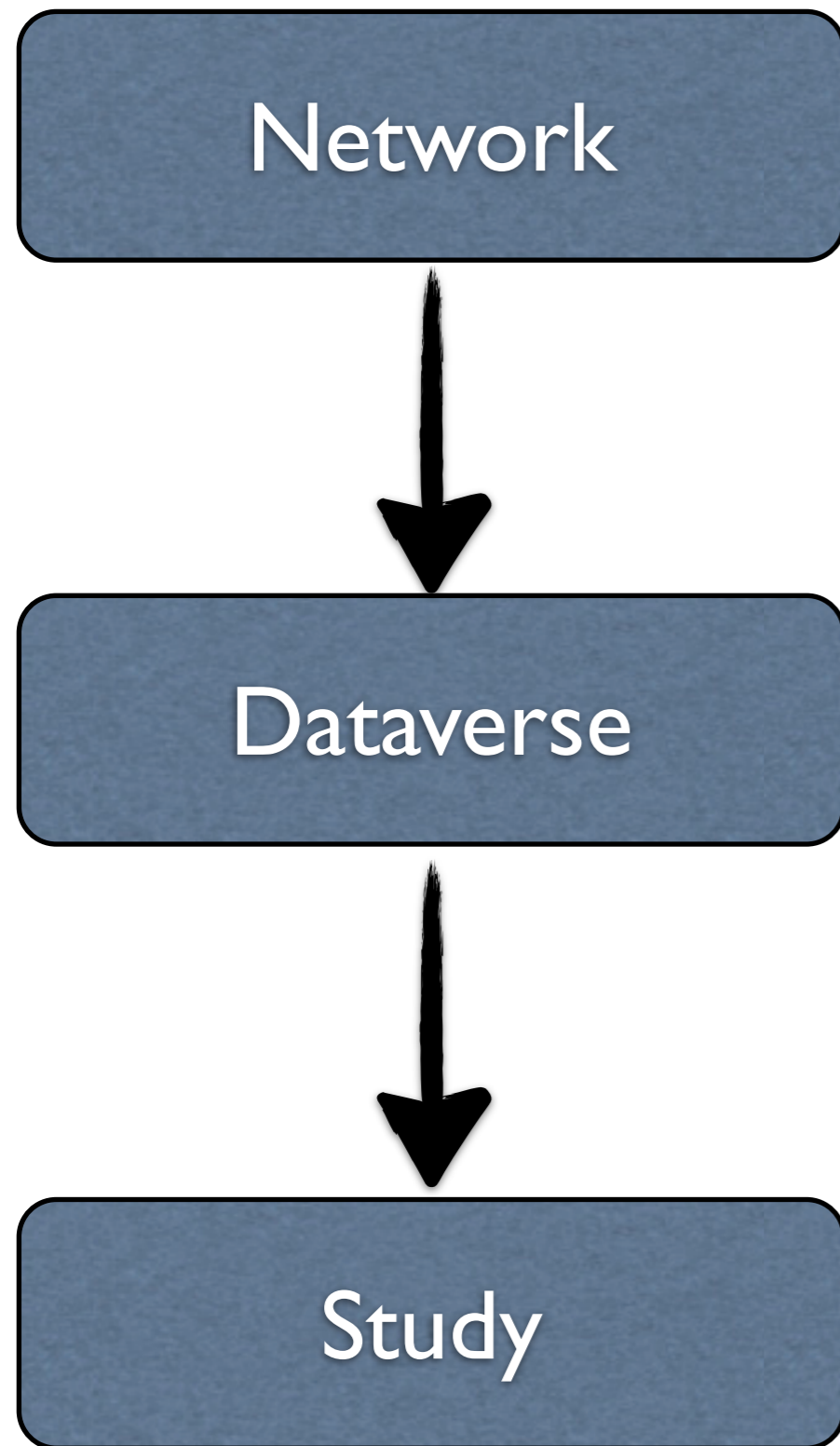
<http://theastrodata.org>



- The Dataverse Network (DVN) Project was built originally for managing Social Science Data;
- Collaboration between the Harvard/CfA “*Seamless Astronomy*” team and the DVN team to reuse this framework for Astronomy Data.
- Institutional support from Harvard Library for DVN infrastructure and training for Astronomy.



- Gives **ownership and recognition** to data owner
- Generates a **persistent data citation**
- Converts data sets to a **preservable** and verifiable format
- Distributes data to the **public**, but also supports **restricted** access
- Indexes all metadata for quick data **discovery**
- Supports **subsetting and analysis** for (some) data files
- Can be branded as your web site.
- Inter-operates** with other systems using **standards**







## *We are:*

- Metadata mapping between the Data Documentation Initiative (DDI) standard used by DVN and Astronomy's VO standards;
- Conducting Data "Interviews" with Astronomers to deduce their needs;
- Working with NASA-SAO ADS to expose data publications;
- Professional Outreach Training for CfA astronomers to use platform;
- Working on the DVN API for search & up/downloading of data products;
- Working with VAO to expose internal data products to VO indexing and search.

<http://figshare.com> ?



## Why DVN?

- Open Source (Java) Software Stack
- Instantiate new Dataverse Networks:
  - Societal, Publishing, Institutional needs.
  - Copy our CfA work to new Astronomy DVN.
- Built in DVN “Universe” search and linking.



## Why DVN?

- Domain Specific
  - Metadata/Data Formats;
  - Use Astronomy Controlled Vocabularies for Curation;
  - Hook up DVN to VO and other Software tools.
  - Reuse DVN API for Astronomy specific software tools



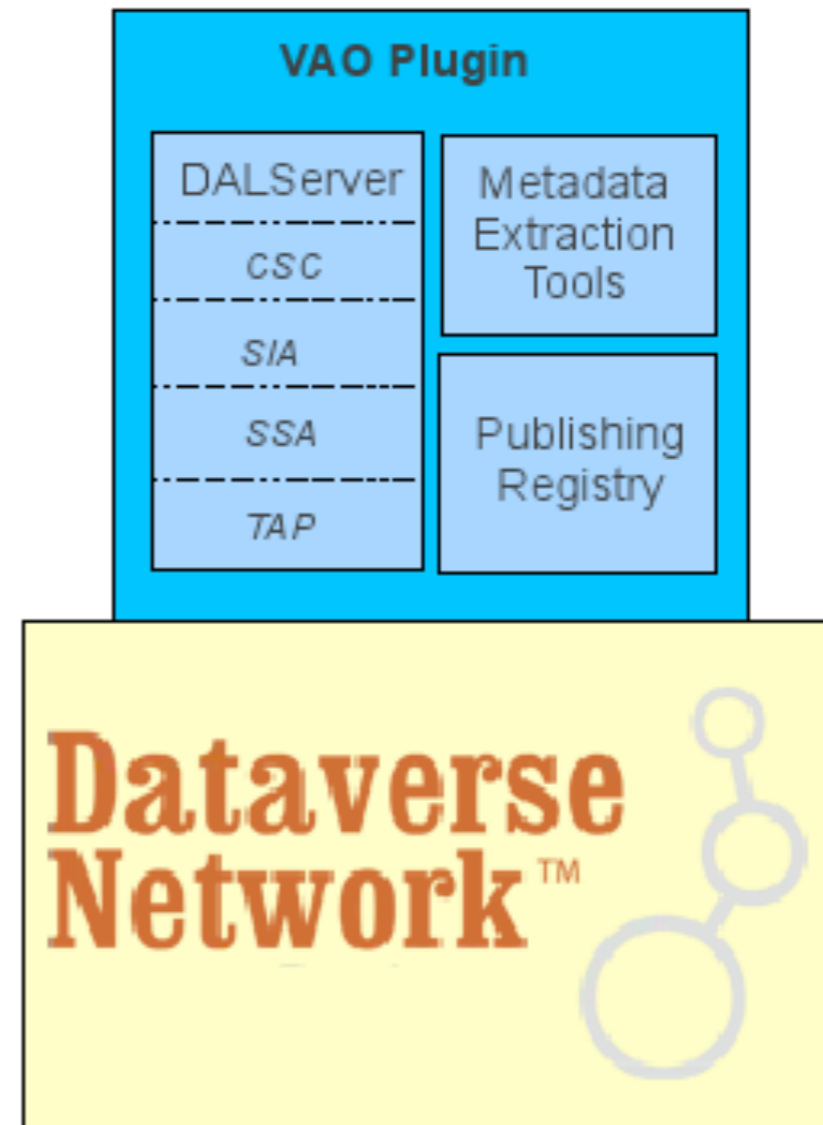
## Why DVN?

- Friends
  - Work with DVN developers to evolve software:
    - Metadata/Data format support.
  - Link Dataverse “Studies”
    - NASA-ADS
    - American Astronomical Society Publications (ApJ, AJ...)



# Virtual Observatory “Plugin” to DVN

- Index individual “datatypes” in a published data study;
- Expose services for datatypes;
- Manage publication registration to VO.





this problem

**References:** Ton Zijlstra; <http://www.flickr.com/photos/tonz/2463875144/>

<http://theastrodata.org>



