Building a community of

TECH SAVVY ASTRONOMERS in the era of big-data and data science

Dr. Arna Karick

Freelance astronomer /e-Research + scientific computing consultant

Based at Inspire9, Melbourne, Australia @drarnakarick www.arnakarick.com | techsavvyastronomer.io

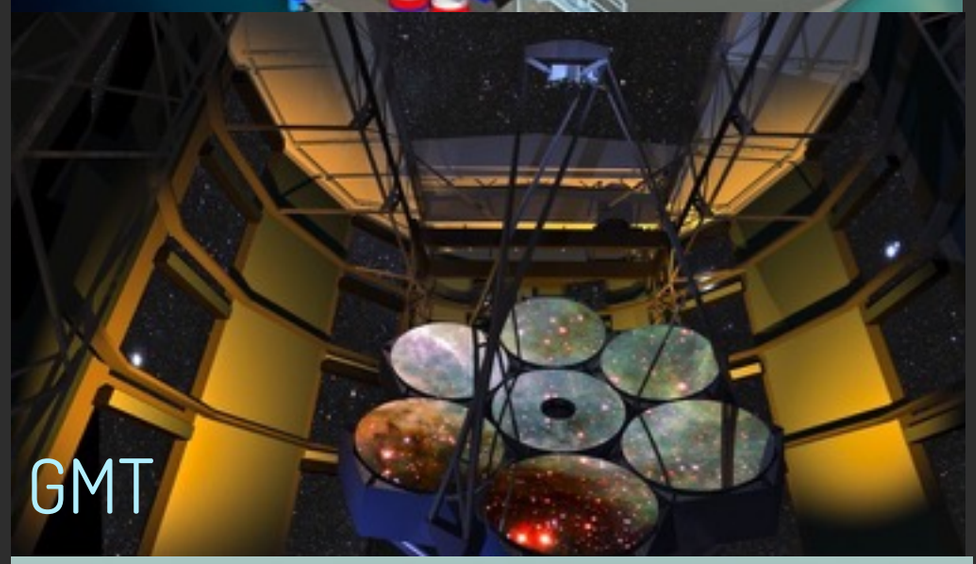
This talk is about building a community of tech savvy astronomers

tech savvy astronomer

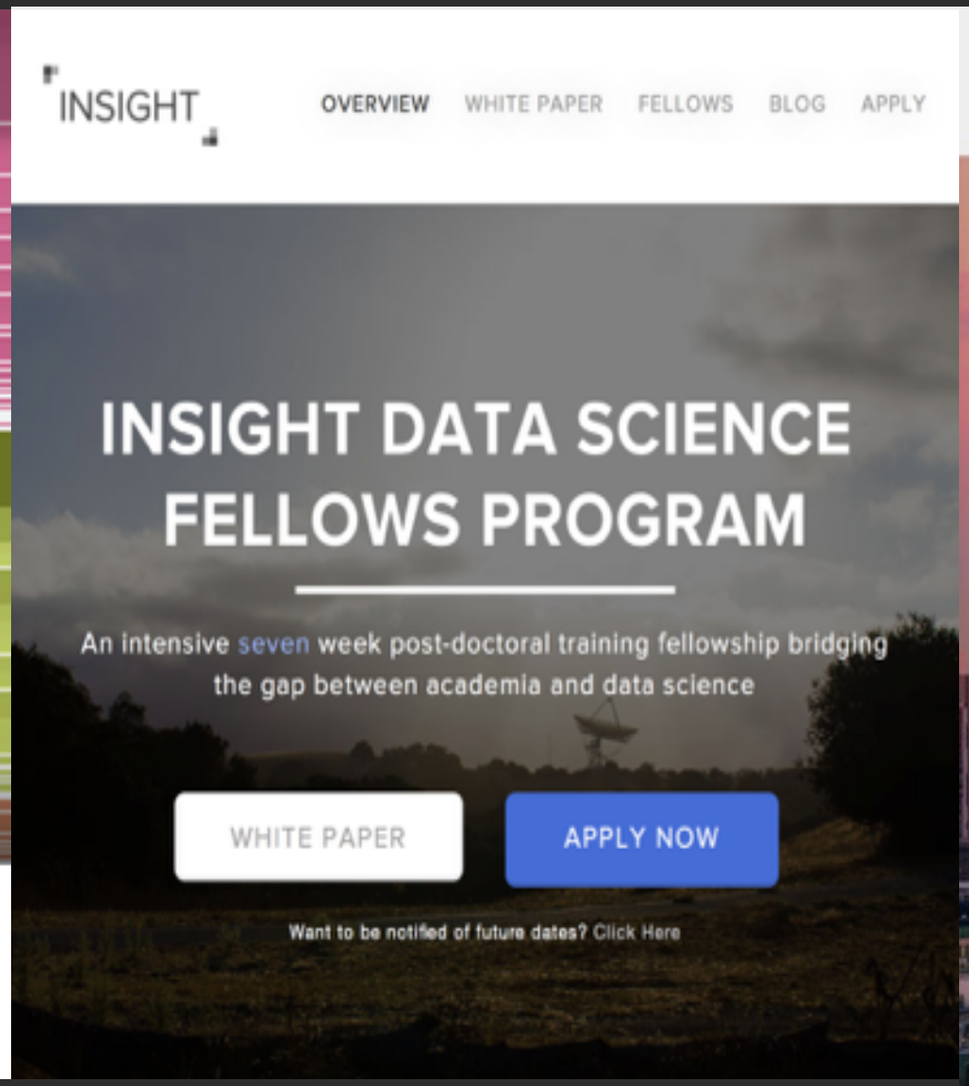
(noun)

- a researcher with tech-focussed and/or ad-hoc development skills in addition to their astronomy specific data analysis skills

a set of skills that enable them to effectively manage the complex (e.g. 3d + temporal), “big” datasets anticipated from next generation telescopes & science-surveys; the ability to contribute to open software and tools development; the ability to build simple tools for their research and broader community; have experience using tools commonly used in the tech industry.



The rise of "data science" has created a generation of astronomers who want to be tech savvy. The Insight and S2DS fellowships facilitate transitioning to the tech industry, but these are competitive. Being a tech savvy astronomer is advantageous.



The US and Europe are leading the charge with various grassroots and data science initiatives.
Australia is slowly catching up.

.astronomy

SPACE TELESCOPE SCIENCE INSTITUTE

GORDON AND BETTY MOORE FOUNDATION

eScience Institute
ADVANCING DATA-INTENSIVE DISCOVERY IN ALL FIELDS

BERKELEY Institute for Data Science

> {SciCoder;}

ANITA
Australian National Institute for Theoretical Astrophysics

The .Astronomy, Astro Hack Week, Python in Astronomy, SPIE/NAM/AAS Hack Days, bring together a diverse community of astronomers – at all levels, instrument scientists, software developers, data wranglers, data scientists, educators, and science communicators.

They provide a forum for discussing best practices in scientific computing, skills sharing, and an opportunity for astronomers to create innovative research and outreach tools in a safe* and collaborative environment.

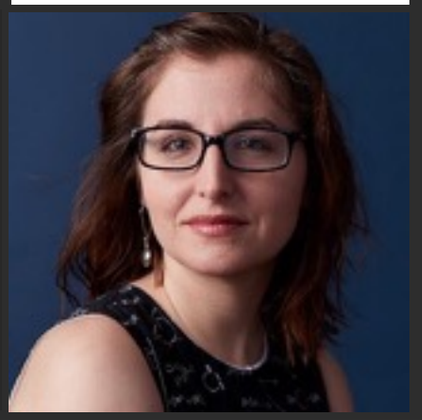
They are participant driven.

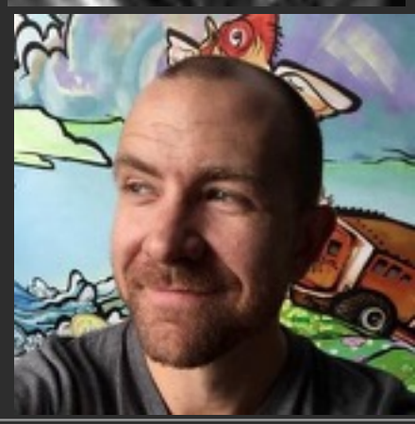
*Imposter syndrome is rife

Previously, astronomers who moved into tech were often lost forever. Those who were a part of the .Astronomy & Astro Hack Week communities continue to be actively involved. Others have become data science mentors.

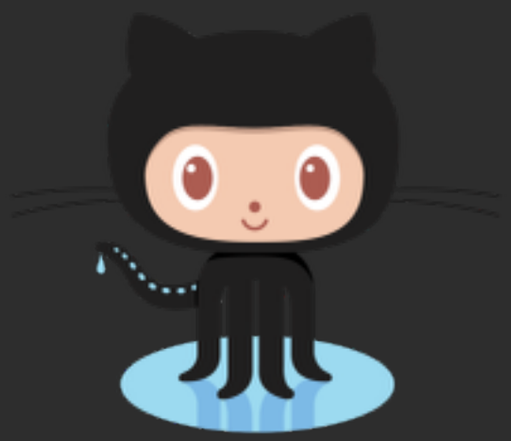
JOURNAL OF OPEN SOURCE SOFTWARE

HIRED





The Journal of Open Source Software: http://joss.theoj.org/



Microsoft Research

GORDON AND BETTY MOORE FOUNDATION

Google

mozilla Science Lab

The Astronomical Society of Australia

Astronomy Australia Ltd.

Microsoft New England Research & Development Center

What actually happens?

Formal talks | Discussions/BoFs | Tutorials | Hacking

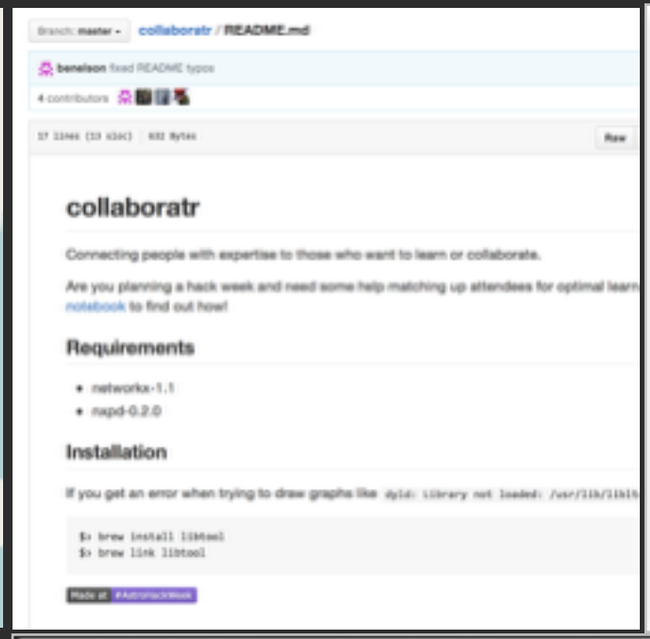
Which typically have this effect on participants...

Tutorials from experts
Software & data publishing: DOIs –> AAS policy
Collaborative coding & source control -> GitHub & BitBucket,
Creating & embedding data visualisations: Aladin Lite,
Interactive data visualisation with D3js & GlueViz
AstroPy & other open development projects
Hacking the literature & reproducible science
Django & Flask web-application frameworks
Building websites, hosting & managing domain names
HTML, CSS & Javascript
Web scraping, using & writing APIs
Mobile applications, web design, wireframes
Sonification of Kepler, IFU, & other multi-wavelength datasets
Visual storytelling & social media hacks (Twitterbots)
code optimisation, machine learning, astrostatistics
Databases: SQLite, DB Browser, SQLAlchemy
Digital Ocean, Docker, IPython, Jupyter Notebooks, Binder, Discourse

At some point magic happens

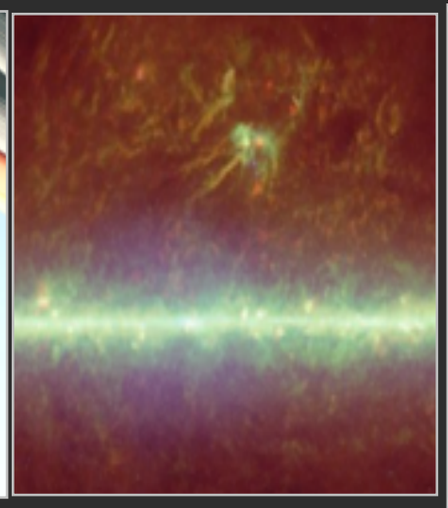

Help astronomers to re-create the incredible stories from the Apollo missions.
Get started!
ABOUT PROJECT APOLLO STORIES
We want to re-create the incredible stories from the Apollo missions, by creating beautiful, interactive timelines. NASA's Project Apollo image archive contains over ~14,000 high resolution images. Help us to pick the most beautiful, inspirational and striking images. The best of the best will go into making interactive timelines and Zooniverse designed image galleries.
Project Apollo Stories Statistics
12
Registered Volunteers
0
Classifications
963
Subjects
0
Retired Subjects

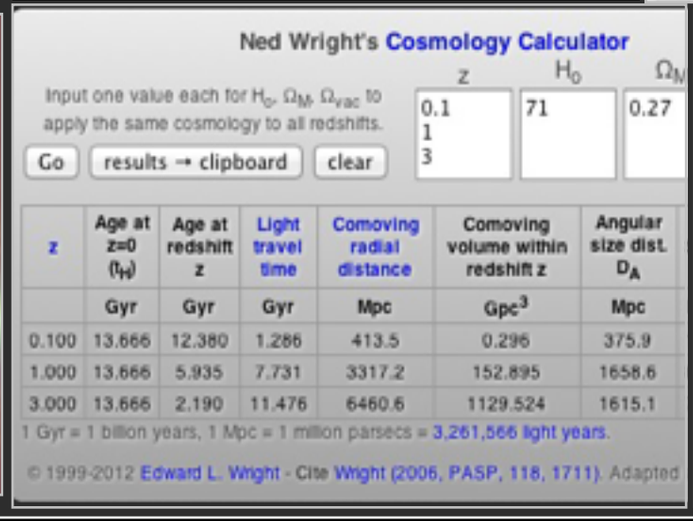

Faculty Positions (visiting and non-tenure)
Faculty Positions (tenure and tenure-track)
Other
Pre-doctoral/Graduate Positions
Post-doctoral Positions and Fellowships
Science Engineering
Science Management
Scientific/Technical Staff

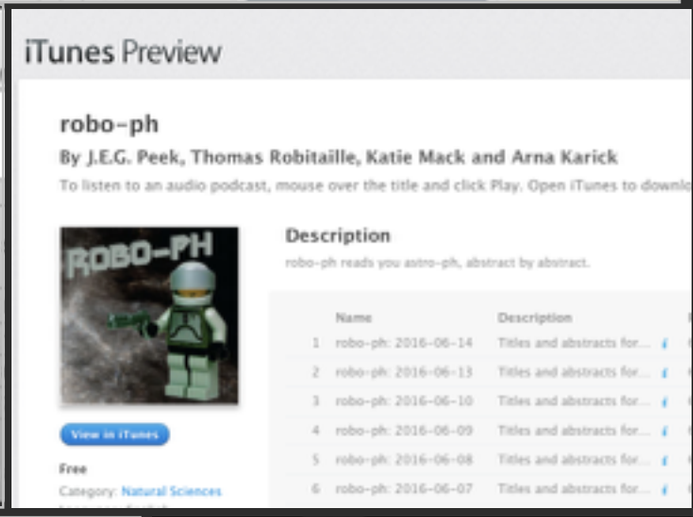

Branch: master
collaboratr / README.md
benelson fixed README typos
4 contributors
collaboratr
Connecting people with expertise to those who want to learn or collaborate.
Are you planning a hack week and need some help matching up attendees for optimal learn
notebook to find out how!
Requirements
networkx-1.1
napd-0.2.0
Installation
If you get an error when trying to draw graphs like dyld: Library not loaded: /usr/lib/libltdl
$> brew install libtool
$> brew link libtool
Made at #AstroHackWeek

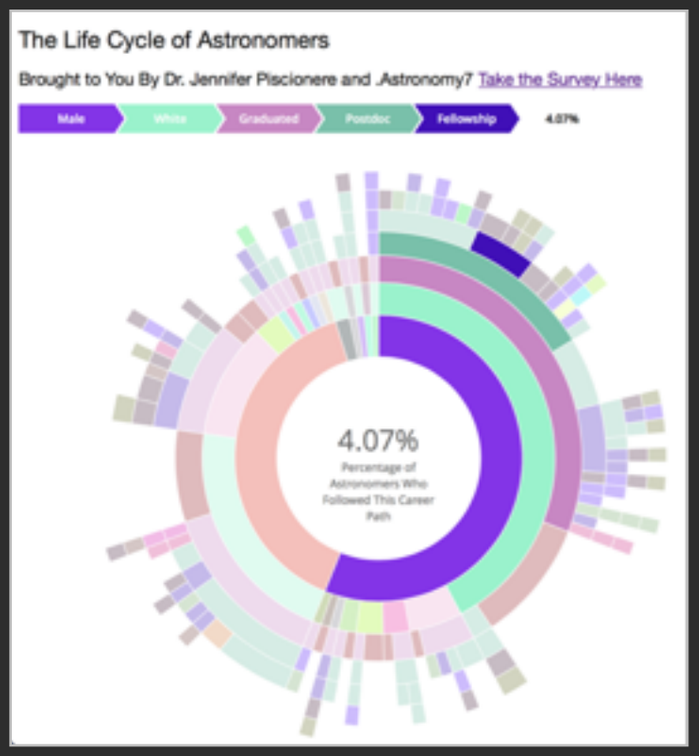

StingRA
Search Tools
OVERVIEW
REVIEWS
RELATED
V* DH Tau


git going with .draft
Write papers on GitHub.
Automagically get PDFs highlighting differences between commits.
Be happy. Make your co-authors happy.

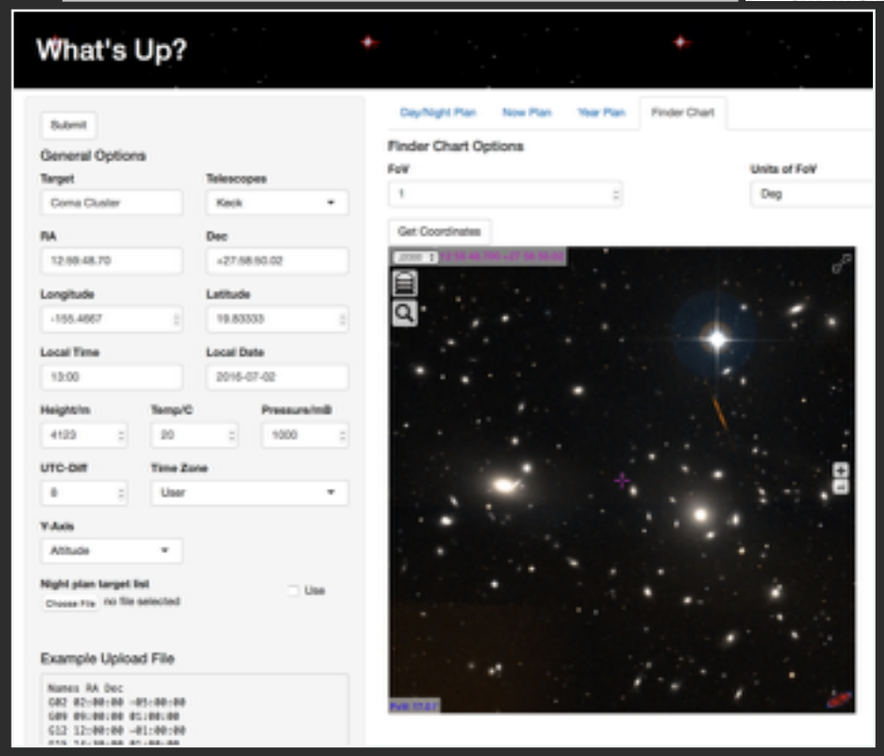


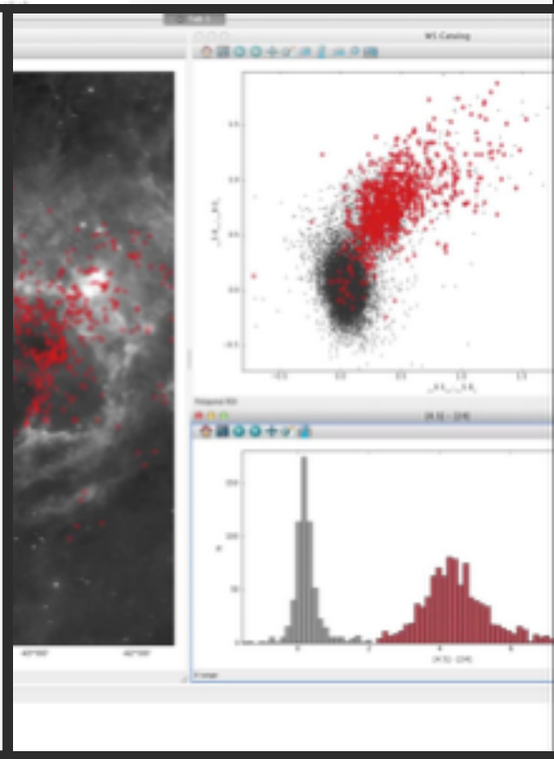

Ned Wright's Cosmology Calculator
Input one value each for H_o, Ω_M, Ω_{vac} to apply the same cosmology to all redshifts.
z
H_0
Ω_M
0.1
1
3
71
0.27
Go
results → clipboard
clear
z
Age at z=0 (t_H)
Age at redshift z
Light travel time
Comoving radial distance
Comoving volume within redshift z
Angular size dist. D_A
Gyr
Gyr
Gyr
Mpc
Gpc^3
Mpc
0.100 13.666 12.380 1.286 413.5 0.296 375.9
1.000 13.666 5.935 7.731 3317.2 152.895 1658.6
3.000 13.666 2.190 11.476 6460.6 1129.524 1615.1
1 Gyr = 1 billion years, 1 Mpc = 1 million parsecs = 3,261,566 light years.
© 1999-2012 Edward L. Wright - Cite Wright (2006, PASP, 118, 1711). Adapted


iTunes Preview
robo-ph
By J.E.G. Peek, Thomas Robitaille, Katie Mack and Arna Karick
To listen to an audio podcast, mouse over the title and click Play. Open iTunes to downlo
ROBO-PH
View in iTunes
Free
Category: Natural Sciences
Description
robo-ph reads you astro-ph, abstract by abstract.
Name
Description
1 robo-ph: 2016-06-14 Titles and abstracts for...
2 robo-ph: 2016-06-13 Titles and abstracts for...
3 robo-ph: 2016-06-10 Titles and abstracts for...
4 robo-ph: 2016-06-09 Titles and abstracts for...
5 robo-ph: 2016-06-08 Titles and abstracts for...
6 robo-ph: 2016-06-07 Titles and abstracts for...


The Life Cycle of Astronomers
Brought to You By Dr. Jennifer Piscionere and .Astronomy7 Take the Survey Here
Male
White
Graduated
Postdoc
Fellowship
4.07%
4.07%
Percentage of Astronomers Who Followed This Career Path


T-Mobile
01:13
PAPR
ironmental dependence of HI
ies in the EAGLE simulations
Cold Streams
Kelvin-
ly Modes
o Marasco, Robert A. Crain, Joop
nick M. Bahé, Thijs van der Hulst,...
hai Dekel,
Mark R. Kru...
Papr
Info


What's Up?
Submit
General Options
Target
Coma Cluster
Telescopes
Keck
RA
12:59:48.70
Dec
+27:58:50.02
Longitude
-155.4667
Latitude
19.83333
Local Time
13:00
Local Date
2016-07-02
Height/m
4123
Temp/C
20
Pressure/mB
1000
UTC-Diff
0
Time Zone
User
Y-Axis
Altitude
Night plan target list
Choose File
No file selected
Use
Example Upload File
Day/Night Plan
Now Plan
Year Plan
Finder Chart
Finder Chart Options
FoV
1
Units of FoV
Deg
Get Coordinates



Unique opportunites: GitHub HQ SF




Astro Hack Week 2016 was offered a whole level of GitHub for one day
Phil Marshall – Stanford: How the LSST DESC uses GitHub for development
Jonathan Whitmore – Silicon Valley Data Science: Jupyter Notebooks
Tour of HQ, dinner, and discussions with GitHub staff

LSST DESC – Development with GitHub: tinyurl.com/how-lsstdesc-uses-github

Value to participants

Expanding your network. Identifying experts.	Learn about tech companies	Tech roles & skills required	Identifying your own abilities & where you can contribute	
Community Developed Tutorials (Jypyter)	Code optimisation	Best practises in scientific computing	Collaborative Coding & Version Control	Introduction to Code Testing
Conversations with software engineers & developers	Statistics/ Baysian Inference	Machine Learning	Getting involved in AstroPy	Building community and a network of experts
Sense of what can be achieved quickly – MVPs	Lean and Agile principles	Code documentation	Learn how languages and applications fit together	Kickstarting new collaborations
Appreciation of real development timelines	What is needed for a robust final product	Complexities of software development	Combining tools to build something entirely new	Discovering useful tools

How does the community benefit?

- Increased collaboration between astronomers, instrument scientists, software developers, engineers, data archivists
- Breaking down barriers between astronomy communities: e.g. ADASS, astroinformatics, .Astronomy, AHW and other grassroots initiatives
- A pool of astronomers able to contribute to the development of data portals, VO projects, software, data analysis & visualisation tools
- Mentors for researchers seeking alternative career paths
- A network of experts within the entire community
- Non-traditional research outputs -> increasing project impact & outreach
- Opportunities for industry engagement and collaboration
- Potential sources of funding, sponsorship
- Start-ups? consultancies? wise.io onekilopars.ec

How do we grow the community?

Skills training programs:

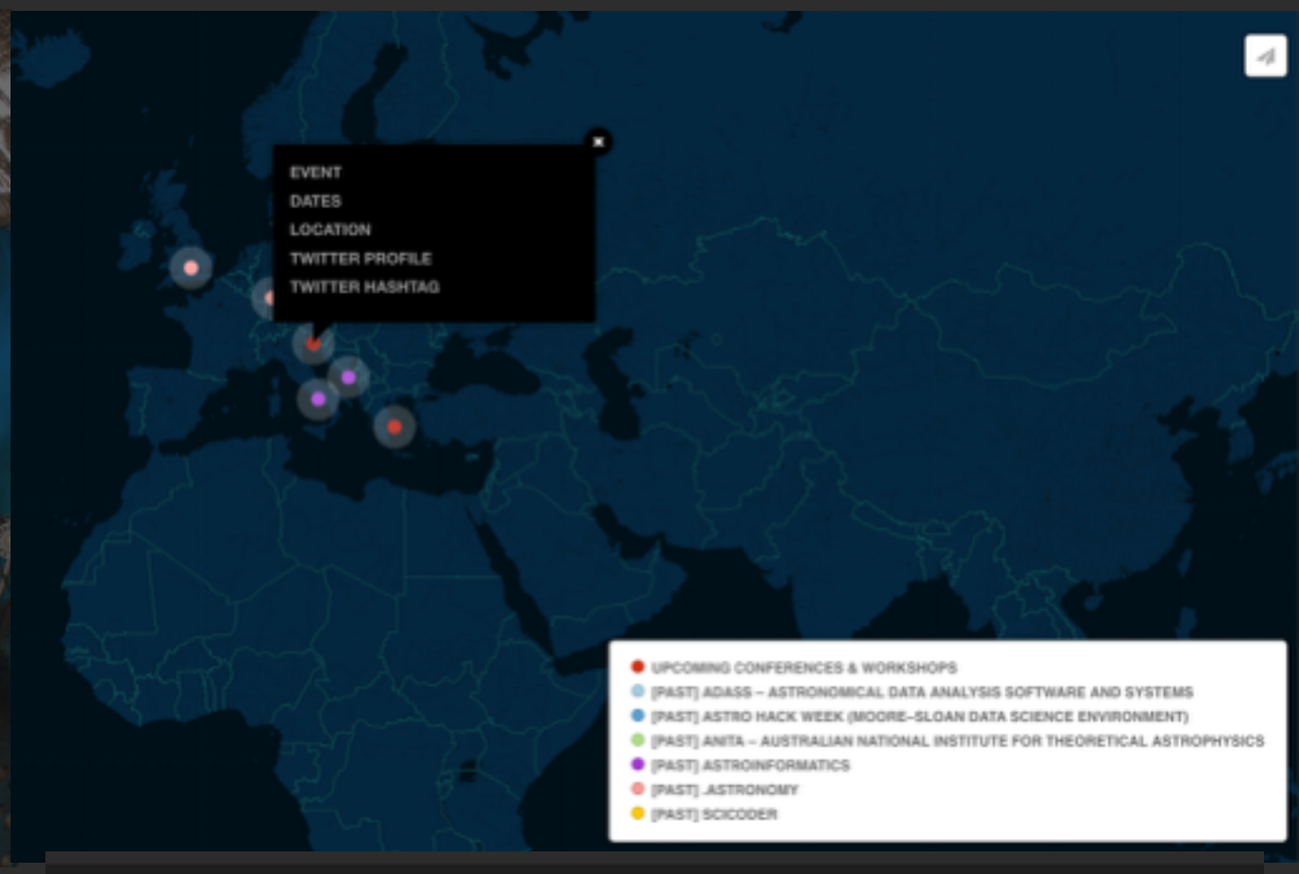
Python, databases, HPC, cloud computing, tech tools, best practice etc.

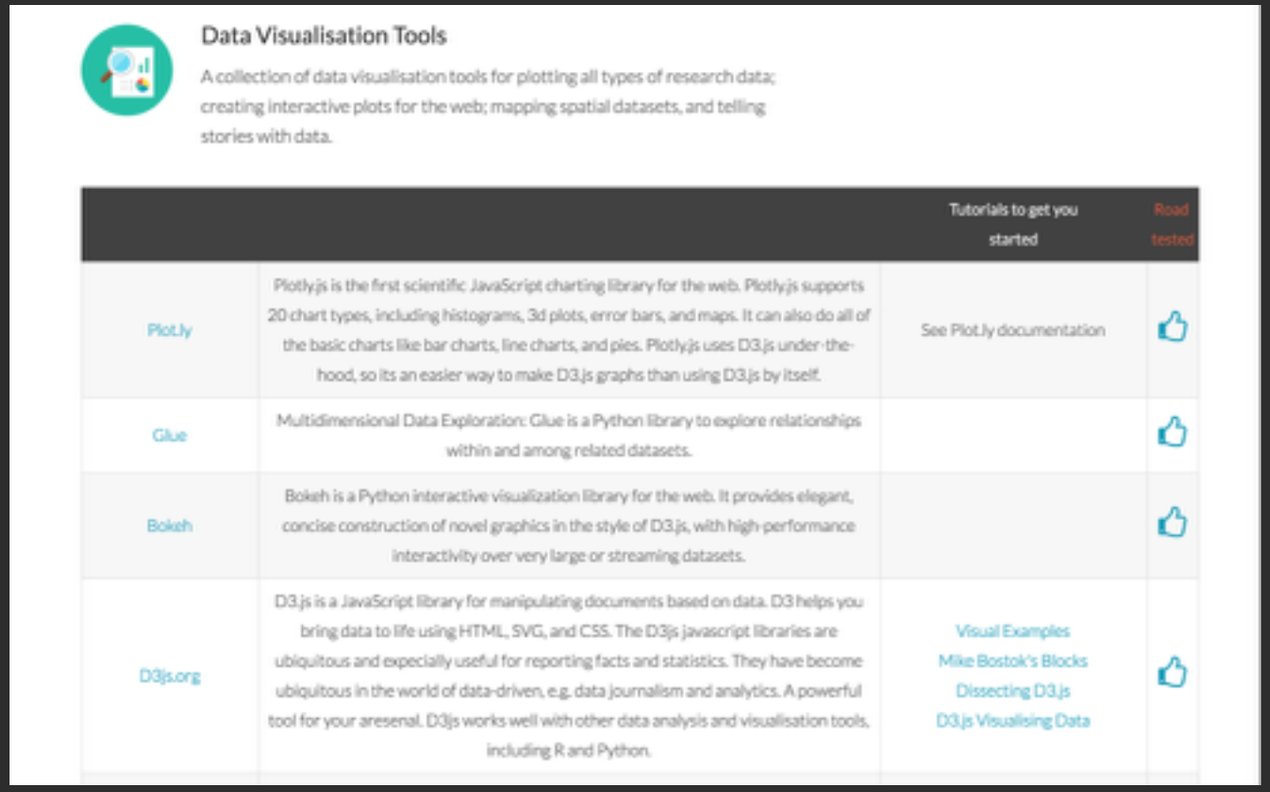
More workshops and events that bring together astronomers at all levels, instrument scientists, software developers & engineers and data scientists.

Hack days focussed on tool building: based around exisiting data portals, early release datasets, or new analysis techniques e.g. machine learning

.Astronomy and Astro Hack Week have benefitted enormously from having software developers, engineers, data archivists and other experts from the ADASS community. Diversity is critical. Get involved.

THANK-YOU

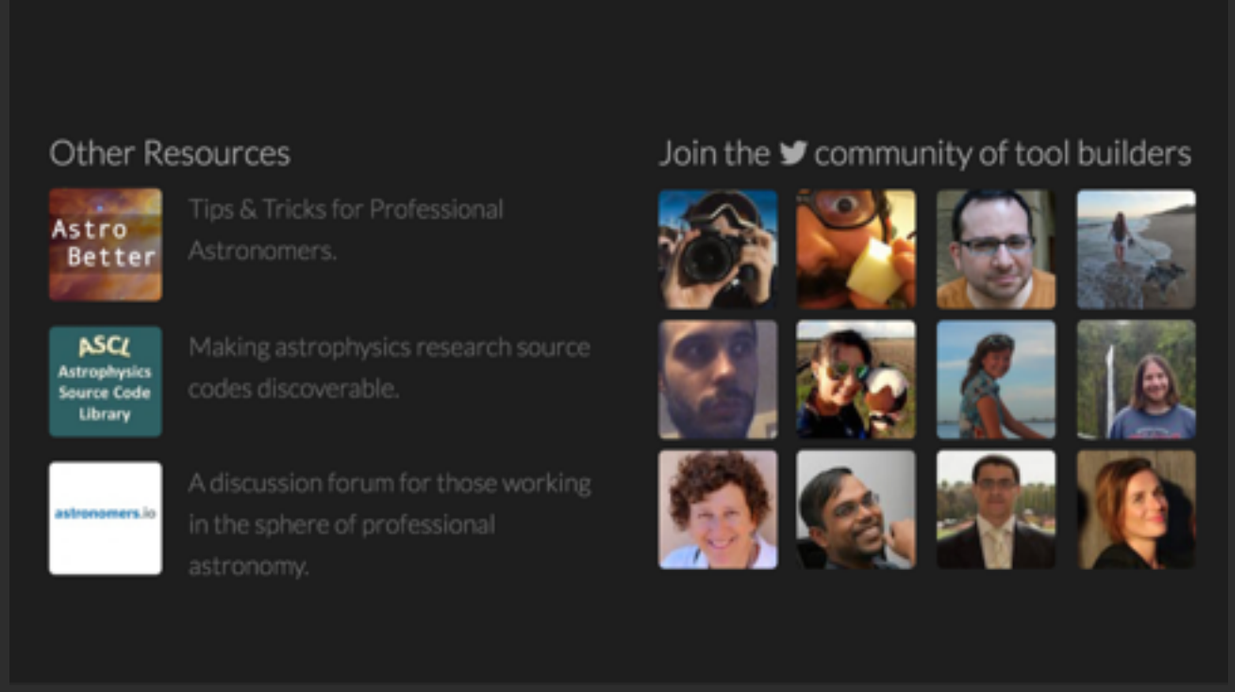




Data Visualisation Tools

A collection of data visualisation tools for plotting all types of research data; creating interactive plots for the web; mapping spatial datasets, and telling stories with data.

		Tutorials to get you started	Road tested
Plot.ly	Plotly.js is the first scientific JavaScript charting library for the web. Plotly.js supports 20 chart types, including histograms, 3d plots, error bars, and maps. It can also do all of the basic charts like bar charts, line charts, and pies. Plotly.js uses D3.js under the hood, so its an easier way to make D3.js graphs than using D3.js by itself.	See Plot.ly documentation	
Glue	Multidimensional Data Exploration: Glue is a Python library to explore relationships within and among related datasets.		
Bokeh	Bokeh is a Python interactive visualization library for the web. It provides elegant, concise construction of novel graphics in the style of D3.js, with high-performance interactivity over very large or streaming datasets.		
D3js.org	D3.js is a JavaScript library for manipulating documents based on data. D3 helps you bring data to life using HTML, SVG, and CSS. The D3js javascript libraries are ubiquitous and expecially useful for reporting facts and statistics. They have become ubiquitous in the world of data-driven, e.g. data journalism and analytics. A powerful tool for your aresenal. D3js works well with other data analysis and visualisation tools, including R and Python.	Visual Examples Mike Bostok's Blocks Dissecting D3.js D3.js Visualising Data	



techsavvyastronomer.io