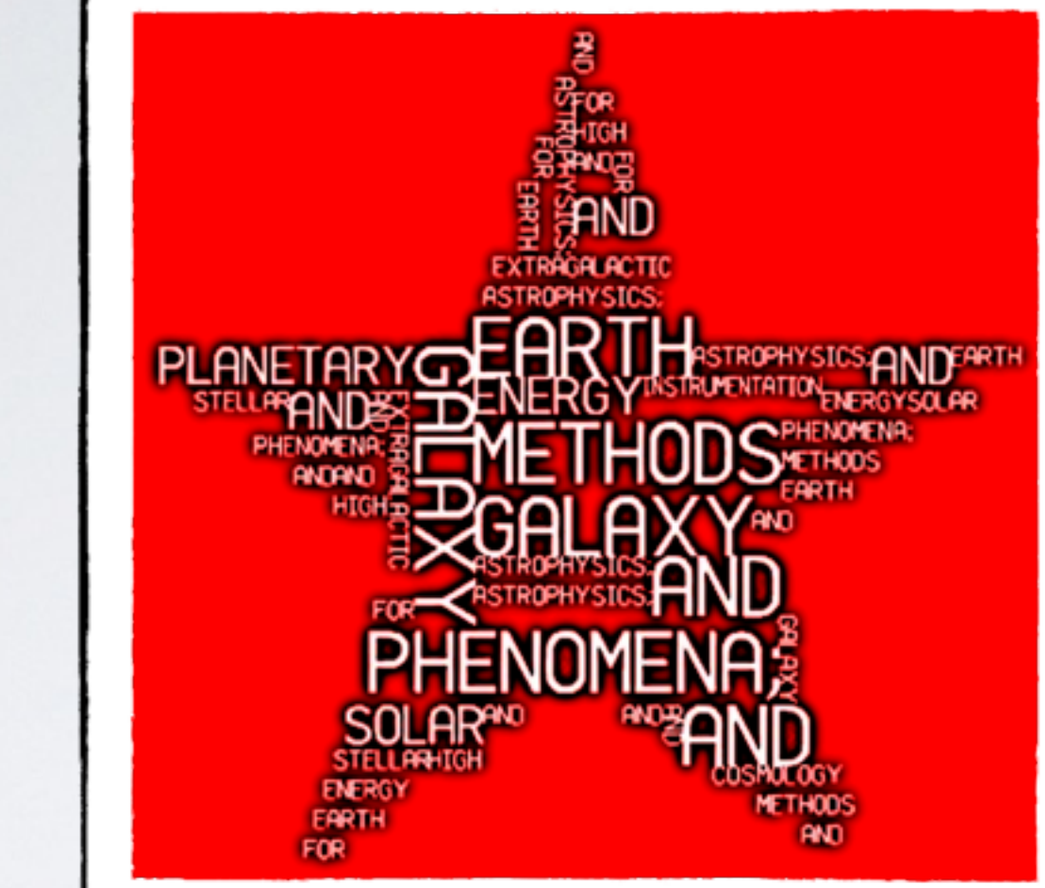


arXiver

Dealing with the big data of scientific literature

Vanessa Moss
CAASTRO/University of Sydney

CAASTRO
ARC CENTRE OF EXCELLENCE FOR ALL-SKY ASTROPHYSICS

Python in Astronomy, 21st April 2015

Using Python as a service

- As knowledgeable programmers, there are various ways that we can use these skills to improve astronomy

- Example: Github repos, software development, education

- These are essentially tools that help people to do astronomy (or anything) in a better/more efficient way

- arXiver is an example of Python as a service, which aims to achieve a similar goal of helping people

My aims through this talk

1. Introduce and explain the workings of arXiver, which was developed using Python to provide a literature service

2. Demonstrate how we have interacted with the community to (hopefully!) improve arXiver over time

3. Give an insight into other ways that Python programmers can use their skills to improve the way we do astronomy

arXiver is brought to you by:

Me & Aidan Hotan (CSIRO)
@cosmicpudding & @EldritchLore

The inner workings of arXiver

Scrape new arXiv postings

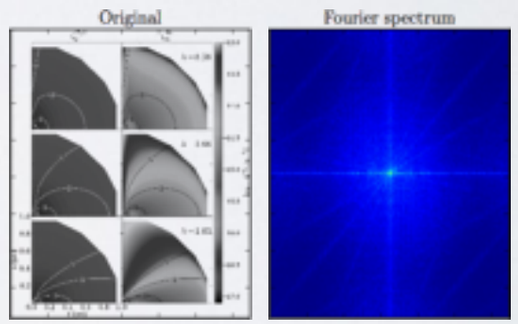
urllib

Download each paper's tar file

urllib

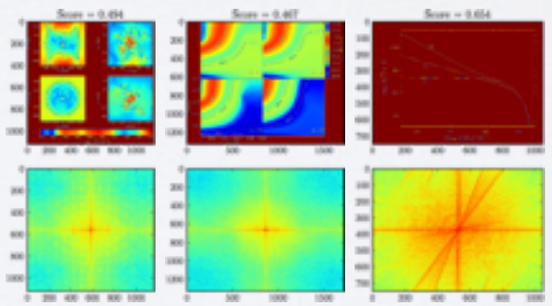
Find associated images

os



Fourier spectral analysis

numpy, scipy



Pick a selection of figures

numpy, scipy

Post to Wordpress.com

wordpress_xmlrpc

Email notification

email, smtplib

The inner workings of arXiver

arXiver daily update: Tue, 21 Apr 15 Inbox x

arxiverbot@gmail.com 09:56 (0 minutes ago)

to me, ahotan, thearxiver

Dear wonderful arXiver monitors,
I have finished posting 69/69 fabulous astro-ph papers for today!
Lots of love,
arXiver

Why is the literature important?

- Science is fundamentally built upon previous work - and astrophysics is no exception to this
- It is (currently) the key quantifiable output of an astrophysicist
- Knowledge of your work's place in a wider field is critical to contextualising its relevance
- Keeping up to date with the newest advances in astronomy gives you an idea of the general direction of the field

How has the literature changed?

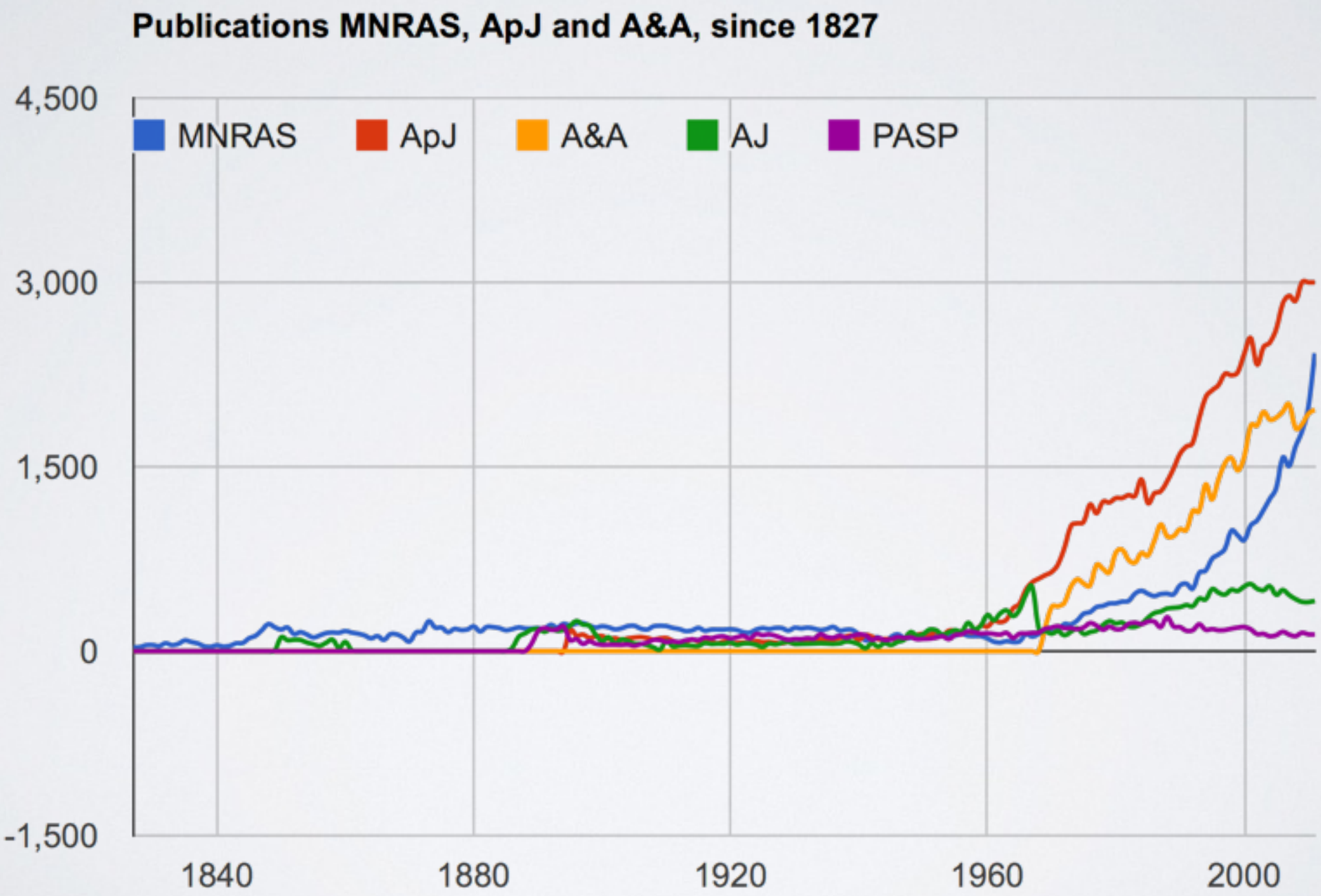
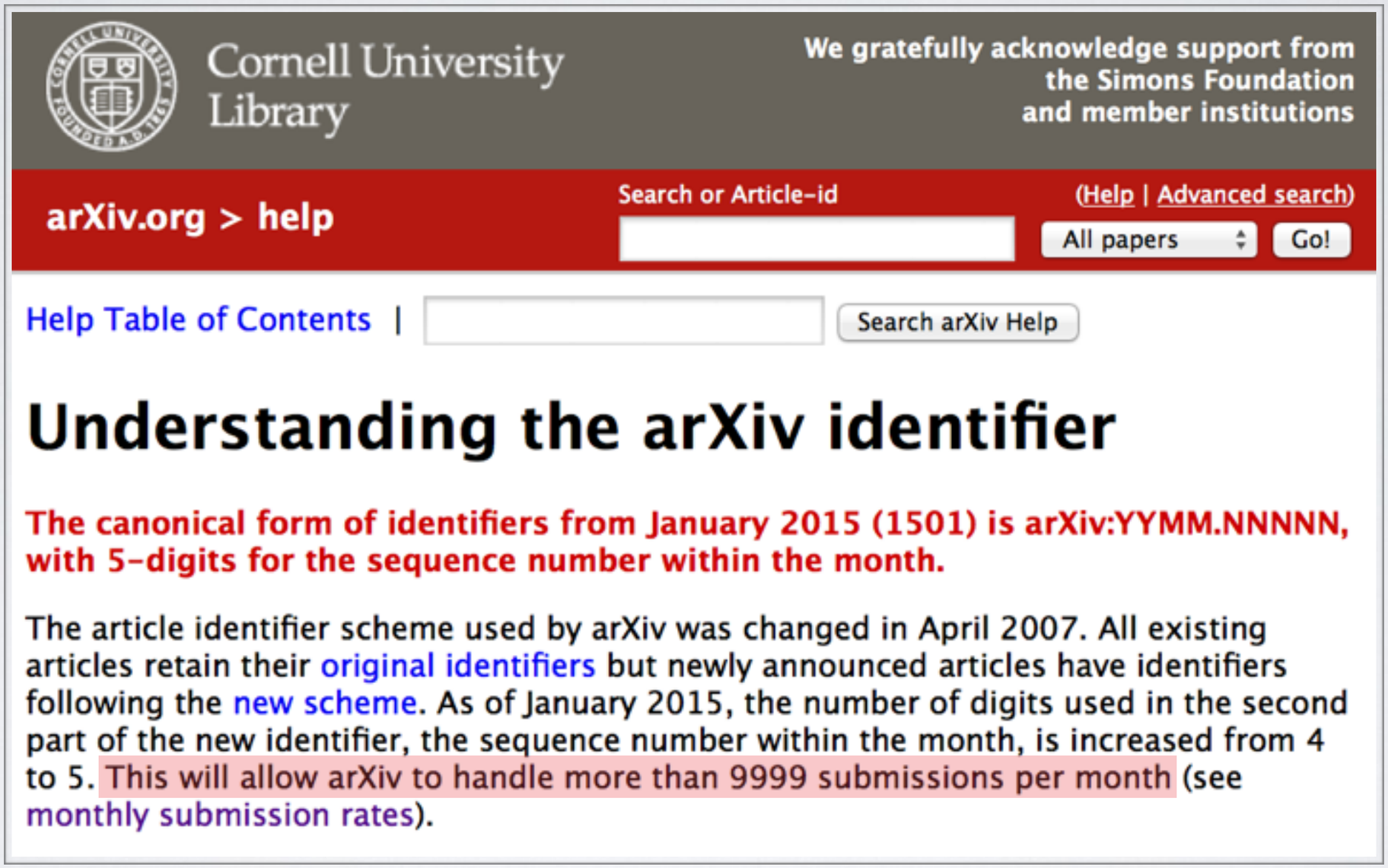


Image credit: Robert Simpson

How has the literature changed?



Cornell University Library

We gratefully acknowledge support from the Simons Foundation and member institutions

arXiv.org > help

Search or Article-id (Help | Advanced search) All papers Go!

Help Table of Contents | Search arXiv Help

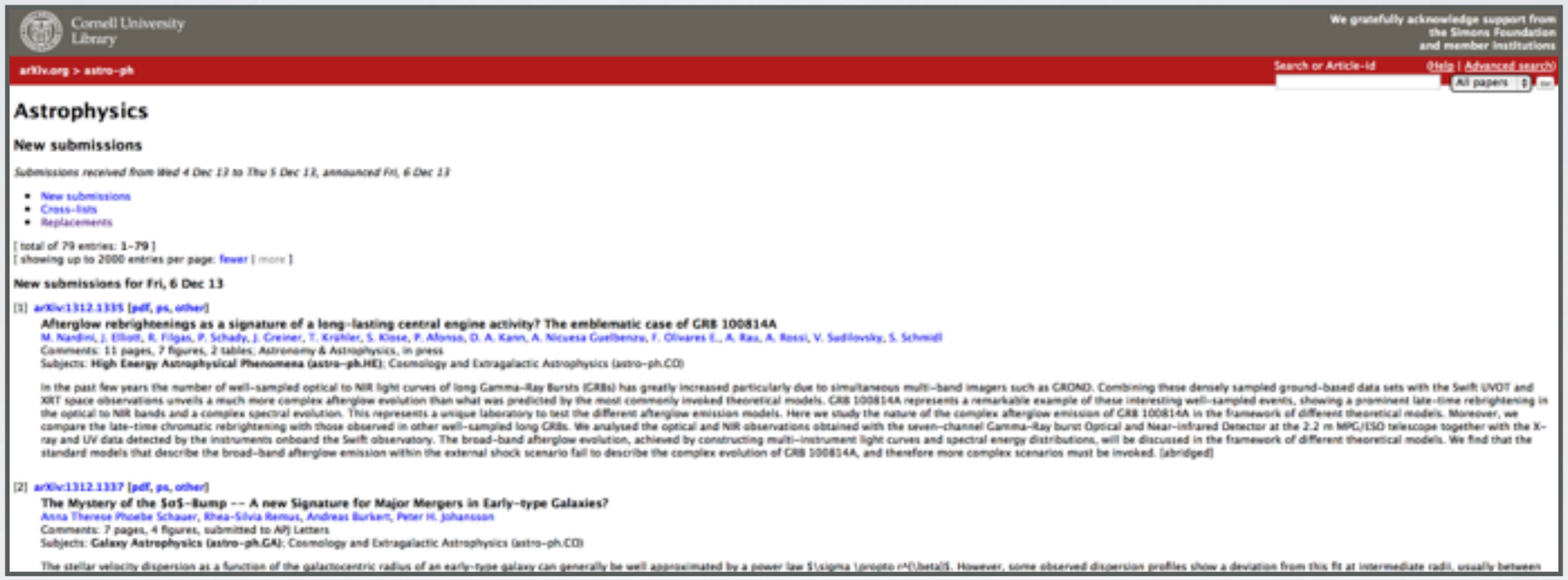
Understanding the arXiv identifier

The canonical form of identifiers from January 2015 (1501) is arXiv:YYMM.NNNNN, with 5-digits for the sequence number within the month.

The article identifier scheme used by arXiv was changed in April 2007. All existing articles retain their original identifiers but newly announced articles have identifiers following the new scheme. As of January 2015, the number of digits used in the second part of the new identifier, the sequence number within the month, is increased from 4 to 5. This will allow arXiv to handle more than 9999 submissions per month (see monthly submission rates).

http://arxiv.org/help/arxiv_identifier

How does arXiver help?



Cornell University Library

We gratefully acknowledge support from the Simons Foundation and member institutions

arXiv.org > astro-ph

Search or Article-id (Help | Advanced search) All papers Go

Astrophysics

New submissions

Submissions received from Wed 4 Dec 13 to Thu 5 Dec 13, announced Fri, 6 Dec 13

- New submissions
- Cross-lists
- Replacements

[total of 79 entries: 1-79]
[showing up to 2000 entries per page: fewer [more]

New submissions for Fri, 6 Dec 13

[1] arXiv:1312.1335 [pdf, ps, other]

Afterglow rebrightenings as a signature of a long-lasting central engine activity? The emblematic case of GRB 100814A

M. Nardini, J. Elliott, R. Filgas, P. Schady, J. Greiner, T. Krühler, S. Klose, P. Afonso, D. A. Kann, A. Nicuesa Guelbenzu, F. Olivares E., A. Rau, A. Rossi, V. Sudilovsky, S. Schmidl

Comments: 11 pages, 7 figures, 2 tables; Astronomy & Astrophysics, in press

Subjects: **High Energy Astrophysical Phenomena (astro-ph.HE)**; Cosmology and Extragalactic Astrophysics (astro-ph.CO)

In the past few years the number of well-sampled optical to NIR light curves of long Gamma-Ray Bursts (GRBs) has greatly increased particularly due to simultaneous multi-band imagers such as GROND. Combining these densely sampled ground-based data sets with the Swift UVOT and XRT space observations unveils a much more complex afterglow evolution than what was predicted by the most commonly invoked theoretical models. GRB 100814A represents a remarkable example of these interesting well-sampled events, showing a prominent late-time rebrightening in the optical to NIR bands and a complex spectral evolution. This represents a unique laboratory to test the different afterglow emission models. Here we study the nature of the complex afterglow emission of GRB 100814A in the framework of different theoretical models. Moreover, we compare the late-time chromatic rebrightening with those observed in other well-sampled long GRBs. We analysed the optical and NIR observations obtained with the seven-channel Gamma-Ray burst Optical and Near-infrared Detector at the 2.2 m MPG/ESO telescope together with the X-ray and UV data detected by the instruments onboard the Swift observatory. The broad-band afterglow evolution, achieved by constructing multi-instrument light curves and spectral energy distributions, will be discussed in the framework of different theoretical models. We find that the standard models that describe the broad-band afterglow emission within the external shock scenario fail to describe the complex evolution of GRB 100814A, and therefore more complex scenarios must be invoked. [abridged]

[2] arXiv:1312.1337 [pdf, ps, other]

The Mystery of the σ-Bump -- A new Signature for Major Mergers in Early-type Galaxies?

Anna Therese Phoebe Schauer, Rhea-Silvia Remus, Andreas Burkert, Peter H. Johansson

Comments: 7 pages, 4 figures, submitted to ApJ Letters

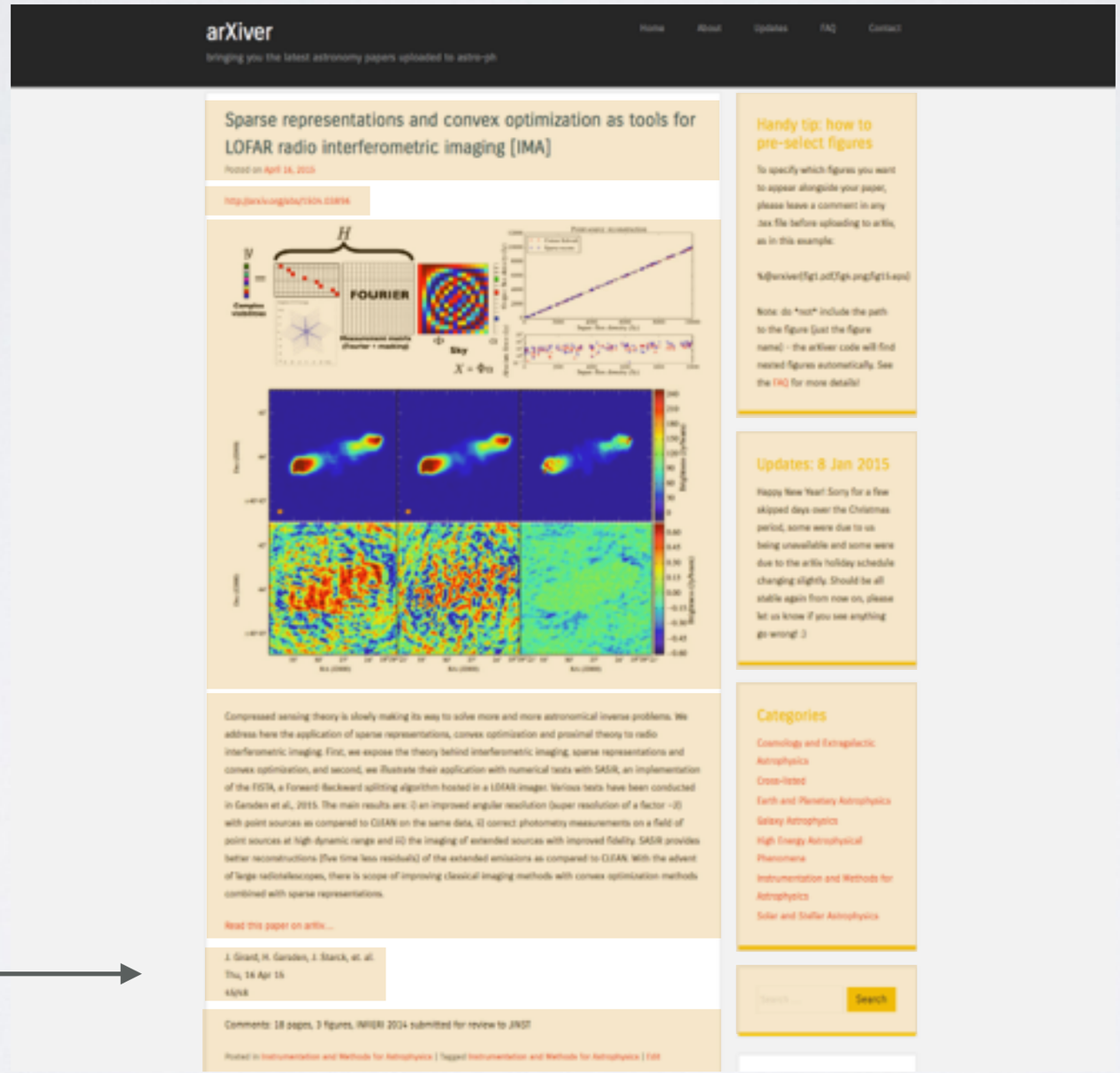
Subjects: **Galaxy Astrophysics (astro-ph.GA)**; Cosmology and Extragalactic Astrophysics (astro-ph.CO)

The stellar velocity dispersion as a function of the galactocentric radius of an early-type galaxy can generally be well approximated by a power law $\sigma \propto r^{\beta}$. However, some observed dispersion profiles show a deviation from this fit at intermediate radii, usually between

- Daily papers appearing on astro-ph can be hard to digest
- We present the meta-data of each paper, reformatted
- We also include three 'representative' figures from the paper

How does arXiver help?

http://arxiver.net



random →

How does arXiver help?

[42] arXiv:1311.5462 [pdf, ps, other]

Hydrodynamical simulations of a compact source scenario for G2

A. Ballone, M. Schartmann, A. Burkert, S. Gillessen, R. Genzel, T.K. Fritz, F. Eisenhauer, O. Pfuhl, T. Ott

Comments: 4 pages, 3 figures; Proceeding of the IAU 303: "The GC: Feeding and Feedback in a Normal Galactic Nucleus" / September 30 – October 4, 2013, Santa Fe, New Mexico (USA)

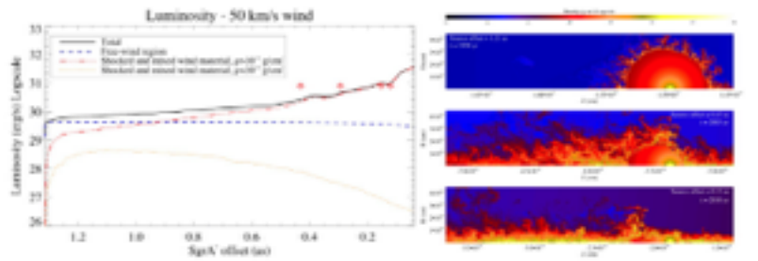
Subjects: **Galaxy Astrophysics (astro-ph.GA)**

The origin of the dense gas cloud G2 discovered in the Galactic Center (Gillessen et al. 2012) is still a debated puzzle. G2 might be a diffuse cloud or the result of an outflow from an invisible star embedded in it. We present here detailed simulations of the evolution of winds on G2's orbit. We find that the hydrodynamic interaction with the hot atmosphere present in the Galactic Center and the extreme gravitational field of the supermassive black hole must be taken in account when modeling such a source scenario. We find that the hydrodynamic interaction with the hot atmosphere present in the Galactic Center and the extreme gravitational field of the supermassive black hole must be taken in account when modeling such a source scenario. We also find that in this scenario most of the Br\gamma\ luminosity is expected to come from the highly filamentary densest shocked wind material. G2's observational properties can be used to constrain the properties of the outflow and our best model has a mass outflow rate of Mdot,w=8.8 x 10^{-8} Msun/yr and a wind velocity of vw = 50 km/s. These values are compatible with those of a young TTauri star wind, as already suggested by Scoville & Burkert (2013).

Hydrodynamical simulations of a compact source scenario for G2 [GA]

Posted on November 22, 2013

http://arxiv.org/abs/1311.5462



The origin of the dense gas cloud G2 discovered in the Galactic Center (Gillessen et al. 2012) is still a debated puzzle. G2 might be a diffuse cloud or the result of an outflow from an invisible star embedded in it. We present here detailed simulations of the evolution of winds on G2's orbit. We find that the hydrodynamic interaction with the hot atmosphere present in the Galactic Center and the extreme gravitational field of the supermassive black hole must be taken in account when modeling such a source scenario. We find that the hydrodynamic interaction with the hot atmosphere present in the Galactic Center and the extreme gravitational field of the supermassive black hole must be taken in account when modeling such a source scenario. We also find that in this scenario most of the Br\gamma\ luminosity is expected to come from the highly filamentary densest shocked wind material. G2's observational properties can be used to constrain the properties of the outflow and our best model has a mass outflow rate of Mdot,w=8.8 x 10^{-8} Msun/yr and a wind velocity of vw = 50 km/s. These values are compatible with those of a young TTauri star wind, as already suggested by Scoville & Burkert (2013).

Read this paper on arXiv...

Fri, 22 Nov 13

58/66

Posted in Galaxy Astrophysics | Tagged Galaxy Astrophysics | Edit

How does arXiver help?

[53] arXiv:1311.5520 [pdf, ps, other]

A Multi-wavelength Study of the Host Environment of SMBHB 4C+37.11

Roger W. Romani, W.R. Forman, Christine Jones, S. S. Murray, A. C. Readhead, Greg B. Taylor

Comments: To appear in the Astrophysical Journal

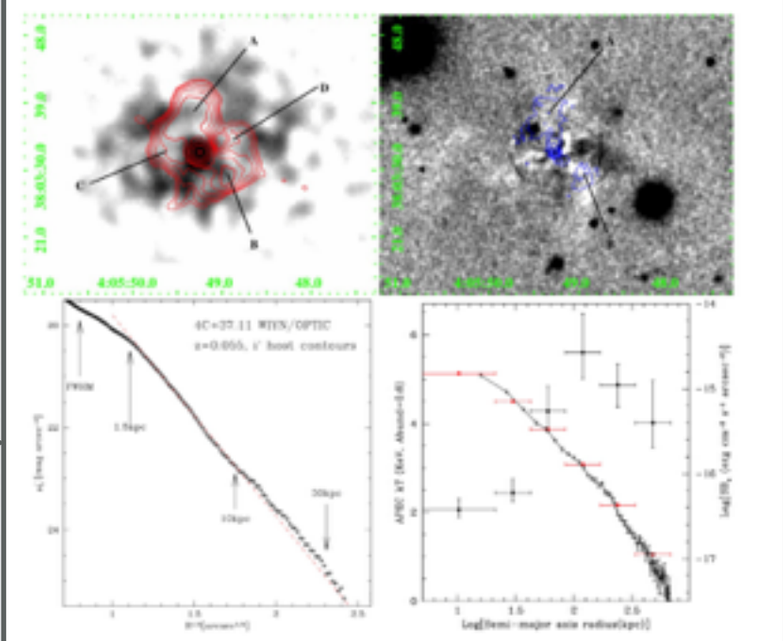
Subjects: **Galaxy Astrophysics (astro-ph.GA)**

4C+37.11, at z=0.055 shows two compact radio nuclei, imaged by VLBI at 7mas separation, making it the closest known resolved super-massive black hole binary (SMBHB). An important question is whether this unique object is young, caught on the way to a gravitational in-spiral and merger, or has `stalled' at 7pc. We describe new radio/optical/X-ray observations of the massive host and its surrounding X-ray halo. These data reveal X-ray/optical channels following the radio outflow and large scale edges in the X-ray halo. These structures are promising targets for further study which should elucidate their relationship to the unique SMBHB core.

A Multi-wavelength Study of the Host Environment of SMBHB 4C+37.11 [GA]

Posted on November 22, 2013

http://arxiv.org/abs/1311.5520



4C+37.11, at z=0.055 shows two compact radio nuclei, imaged by VLBI at 7mas separation, making it the closest known resolved super-massive black hole binary (SMBHB). An important question is whether this unique object is young, caught on the way to a gravitational in-spiral and merger, or has 'stalled' at 7pc. We describe new radio/optical/X-ray observations of the massive host and its surrounding X-ray halo. These data reveal X-ray/optical channels following the radio outflow and large scale edges in the X-ray halo. These structures are promising targets for further study which should elucidate their relationship to the unique SMBHB core.

Read this paper on arXiv...

Fri, 22 Nov 13
24/66

Posted in Galaxy Astrophysics | Tagged Galaxy Astrophysics | Edit

How does arXiver help?

[25] arXiv:1311.5282 [pdf, other]

RESOLVE: A new algorithm for aperture synthesis imaging of extended emission in radio astronomy

H.Junklewitz, M.R.Bell, M.Selig, T.A.Enßlin

Comments: 22 pages, 8 figures, submitted to Astronomy & Astrophysics

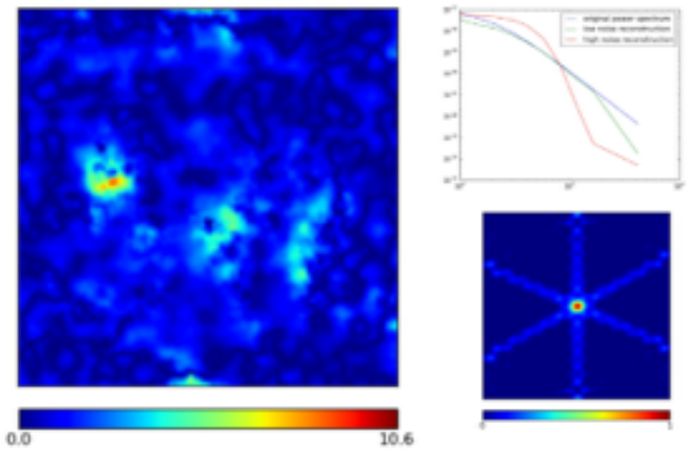
Subjects: **Instrumentation and Methods for Astrophysics (astro-ph.IM)**

We present RESOLVE, a new algorithm for radio aperture synthesis imaging of extended and diffuse emission in total intensity. The algorithm is derived using Bayesian statistical inference techniques, estimating the surface brightness in the sky assuming a priori log-normal statistics. RESOLVE not only estimates the measured sky brightness in total intensity, but also its spatial correlation structure, which is used to guide the algorithm to an optimal reconstruction of extended and diffuse sources. For a radio interferometer, it succeeds in deconvolving the effects of the instrumental point spread function during this process. Additionally, RESOLVE provides a map with an uncertainty estimate of the reconstructed surface brightness. Furthermore, with RESOLVE we introduce a new, optimal visibility weighting scheme that can be viewed as an extension to robust weighting. In tests using simulated observations, the algorithm shows improved performance against two standard imaging approaches for extended sources, Multiscale-CLEAN and the Maximum Entropy Method.

RESOLVE: A new algorithm for aperture synthesis imaging of extended emission in radio astronomy [IMA]

Posted on November 22, 2013

http://arxiv.org/abs/1311.5282



We present RESOLVE, a new algorithm for radio aperture synthesis imaging of extended and diffuse emission in total intensity. The algorithm is derived using Bayesian statistical inference techniques, estimating the surface brightness in the sky assuming a priori log-normal statistics. RESOLVE not only estimates the measured sky brightness in total intensity, but also its spatial correlation structure, which is used to guide the algorithm to an optimal reconstruction of extended and diffuse sources. For a radio interferometer, it succeeds in deconvolving the effects of the instrumental point spread function during this process. Additionally, RESOLVE provides a map with an uncertainty estimate of the reconstructed surface brightness. Furthermore, with RESOLVE we introduce a new, optimal visibility weighting scheme that can be viewed as an extension to robust weighting. In tests using simulated observations, the algorithm shows improved performance against two standard imaging approaches for extended sources, Multiscale-CLEAN and the Maximum Entropy Method.

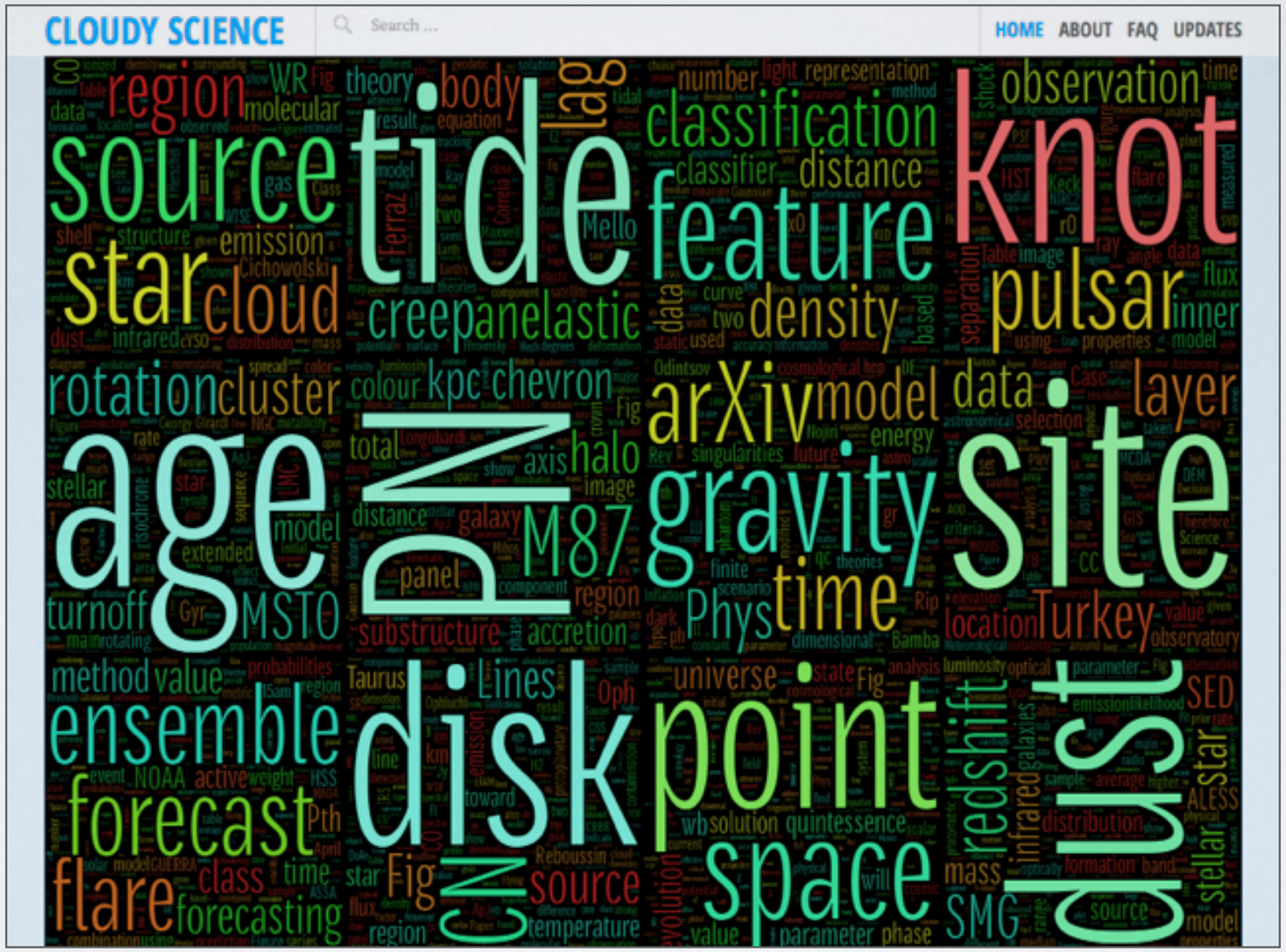
Read this paper on arXiv...

Fri, 22 Nov 13
10/66

Posted in Instrumentation and Methods for Astrophysics | Tagged Instrumentation and Methods for Astrophysics | Edit

“But I prefer words...”

cloudyscience.wordpress.com



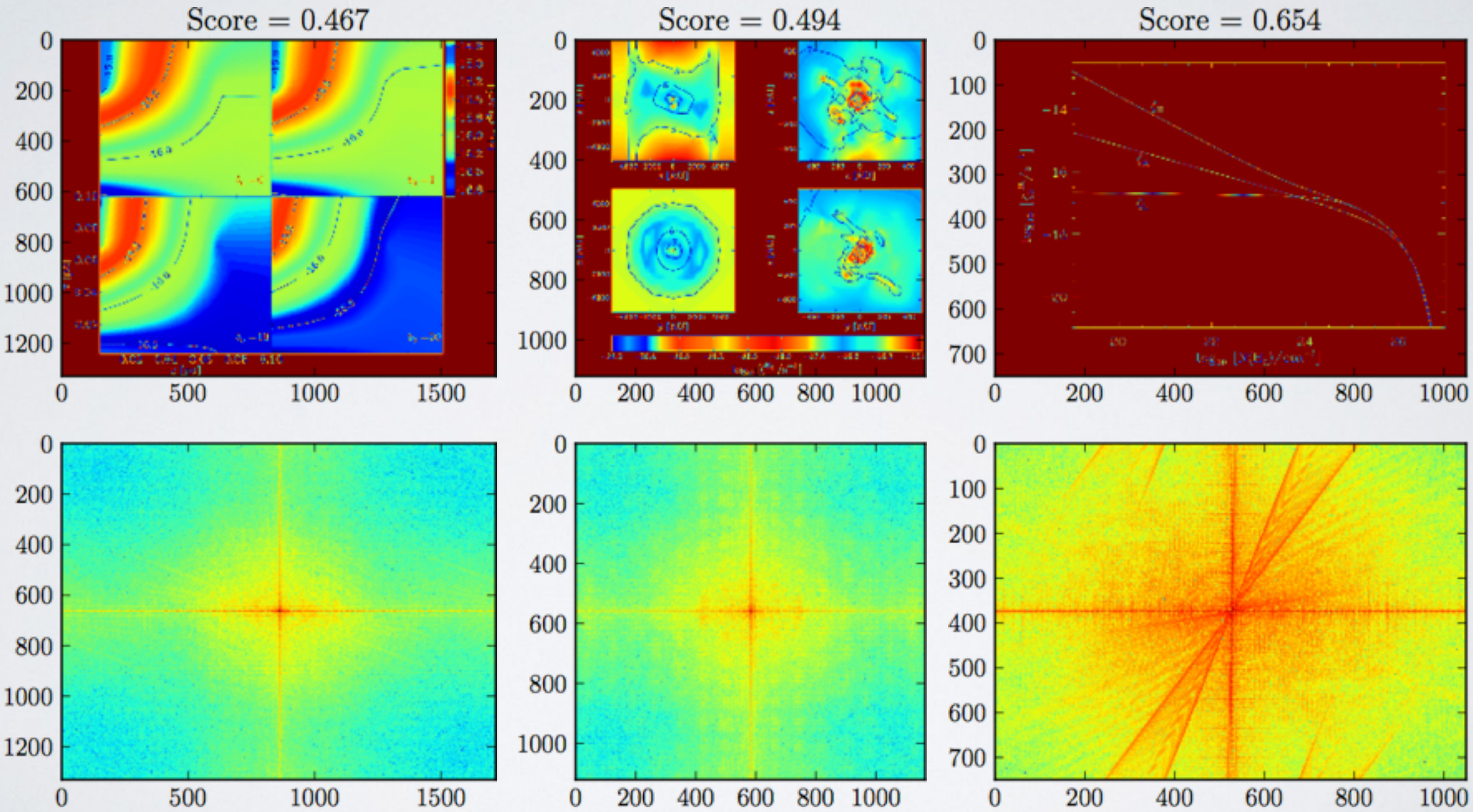
Selecting representative images

- How can an automated process decide what most appeals to a wide variety of human viewers?

- Could apply machine learning with a training data set
  - This seems overly complicated
  - Likely to be high variance in the input data (differing opinions)

- Instead, consider objective information content
  - Scientific information content is impossible to rate without intelligent interpretation, but we can judge the raw complexity of an image
  - A perfect opportunity to apply the Fourier transform!

Making a selection



Initial response to arXiver

Vanessa Moss
arXiver: bringing you the latest astronomy papers uploaded to astro-ph!
http://arxiver.wordpress.com/

We've put together a WordPress that automatically displays the latest papers from astro-ph, together with a few images from each paper in a nice readable format. This developed based off an idea from Rob Simpson about making arXiv easier to follow (and look at!). Each post is also automatically fed to a Twitter account (@arXiver) in case you want to follow it that way.

arXiver has really only been live for a few days and we're still improving various things, but hopefully this will be useful to people in this group 🙂 We'd love your feedback too!

Vanessa Moss and Aidan Hotan

arXiver
arxiver.wordpress.com
bringing you the latest astronomy papers uploaded to astro-ph

Like · Comment · Share · Stop Notifications · October 11 at 11:52am near Sydney

Bryan Gaensler, Billy Robbins, Amanda Bauer and 171 others like this.

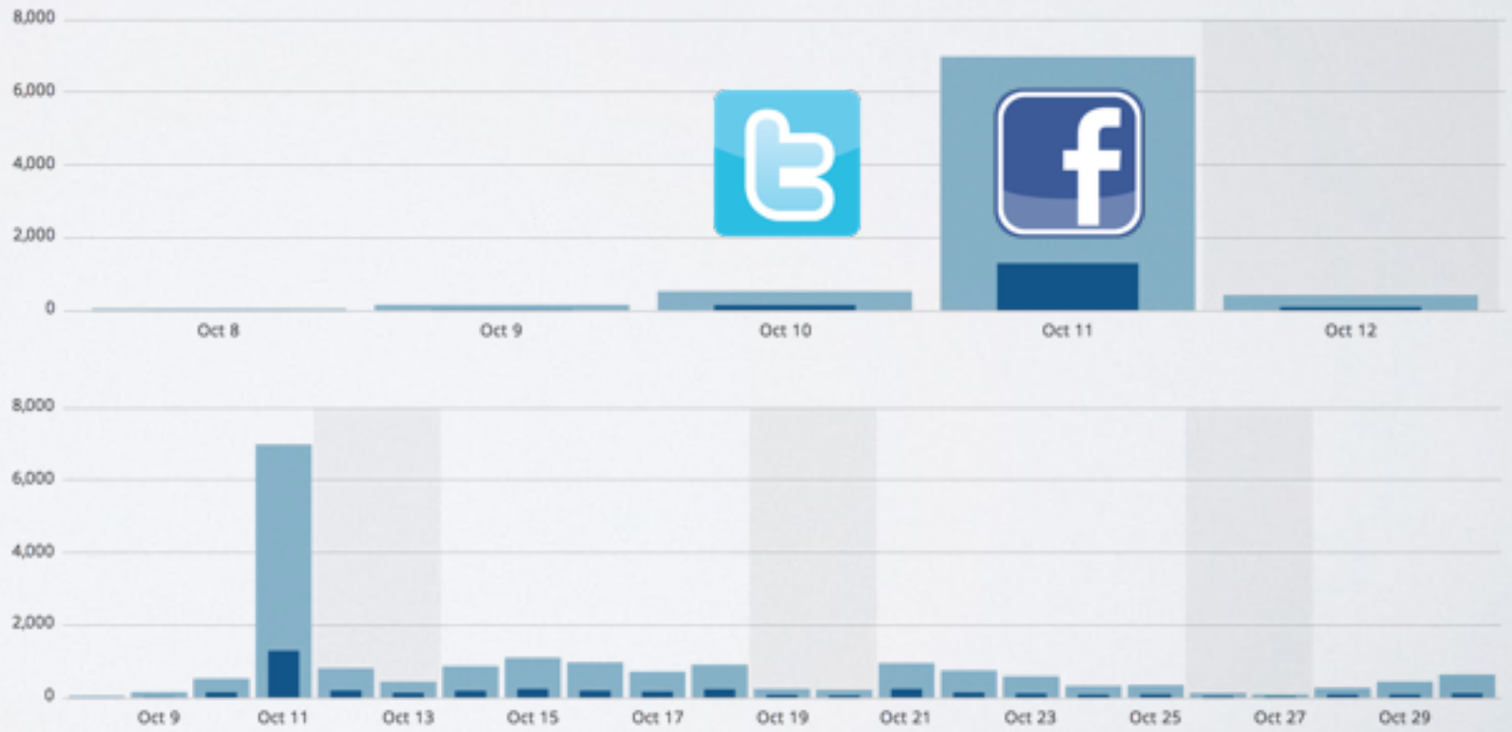
View previous comments 50 of 74

Ivy Wong @owning_ivy 58d
a very nice alternative to the usual arXiv listing thanks to @cosmicpudding & @EldritchLore. About wp.me/P3ZmZy-1 via @arXiver

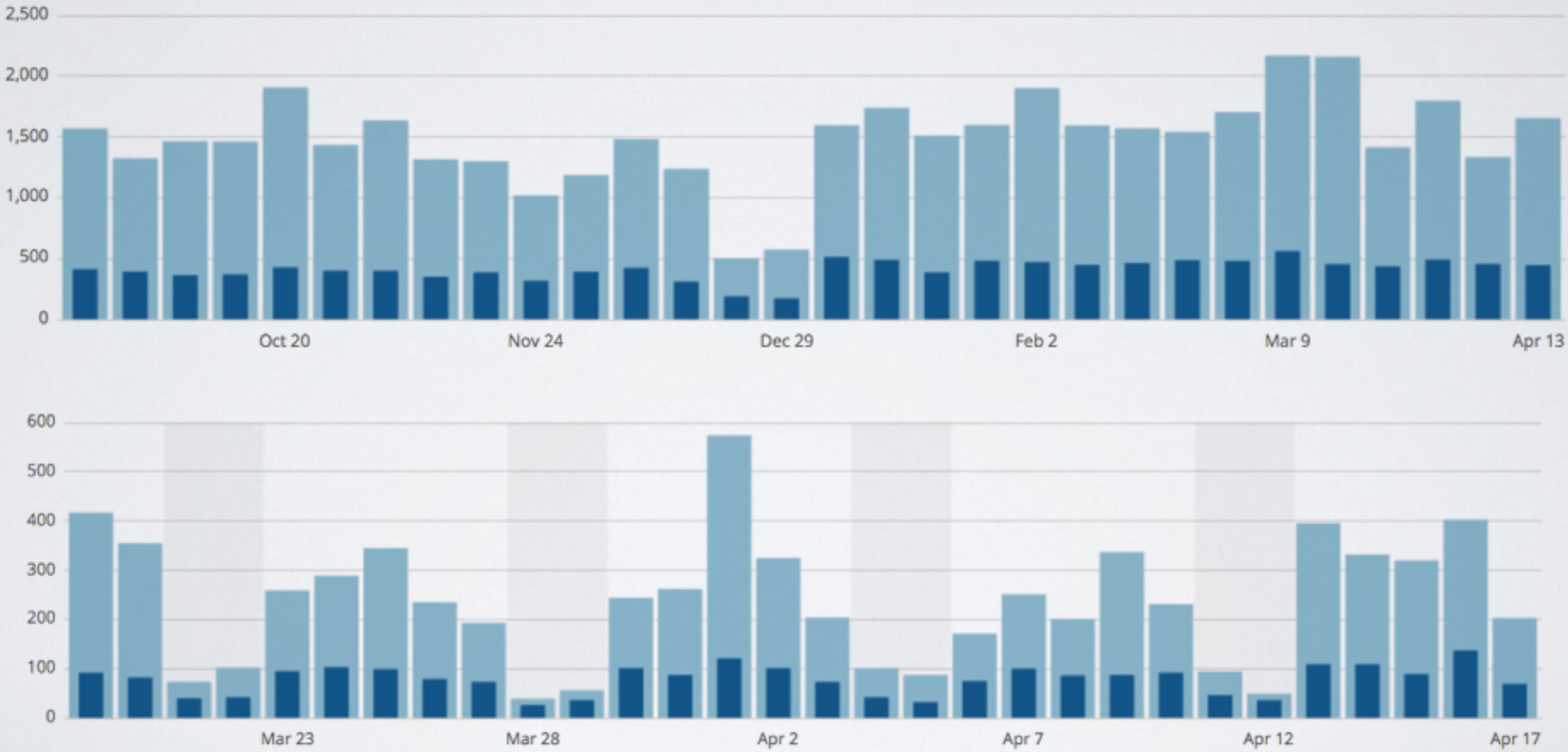
Peter Edmonds @peterdedmonds 57d
A different way to look at new astro papers on the arXiv: arxiver.wordpress.com or @arXiver Give it a try.

★ amanda bauer ★ @astropixie 58d
astronomers, meet @arXiver: a better way to get your daily dose of arXiv! arxiver.wordpress.com/about/ thanks to @cosmicpudding & @EldritchLore

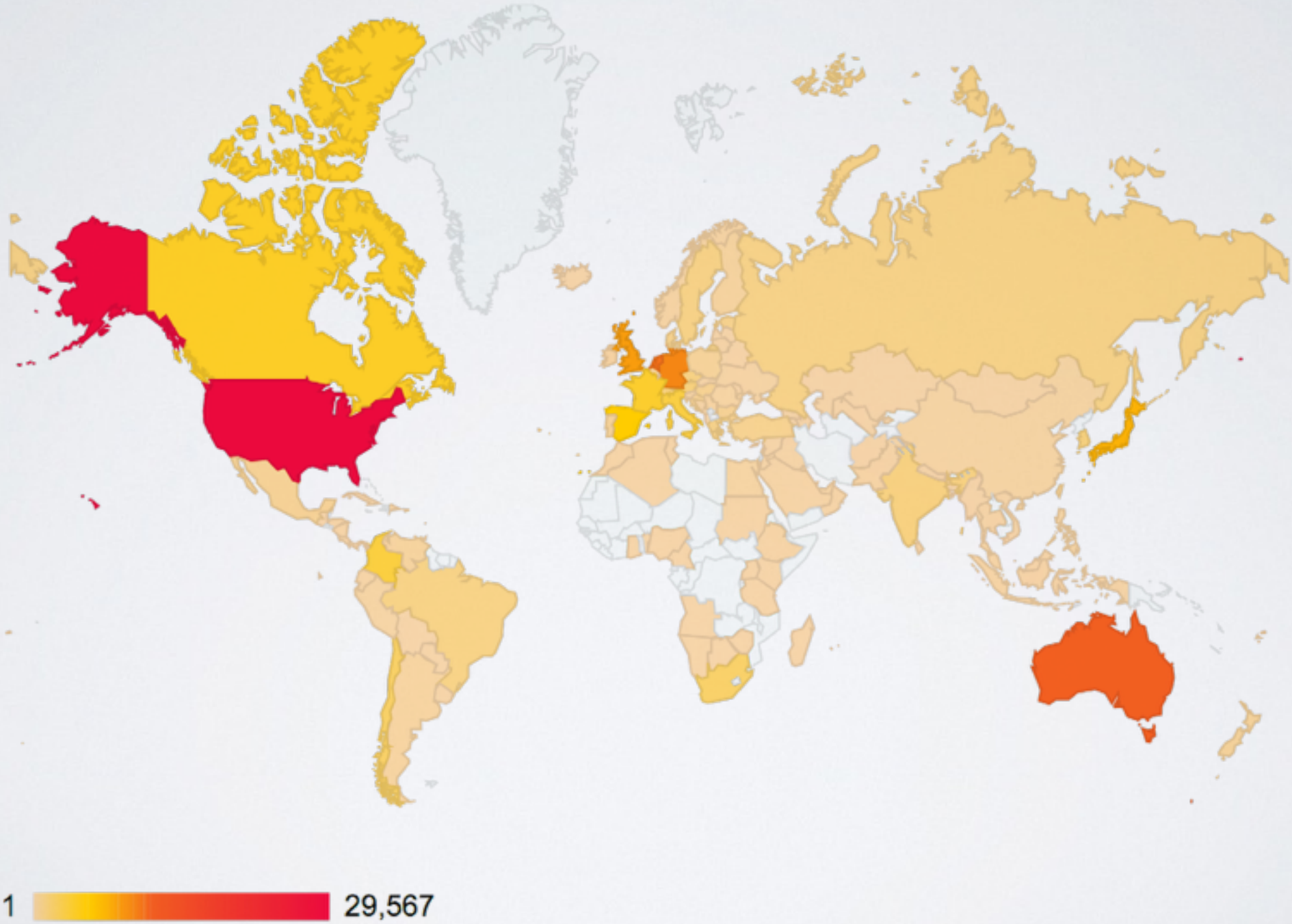
August Muench @augustmuench 57d
Awesome new "figure" based view of arXiv: (@arXiver) bit.ly/18UR9Up -- Hats off to @cosmicpudding & @EldritchLore!



Current usage of arXiver



Current usage of arXiver



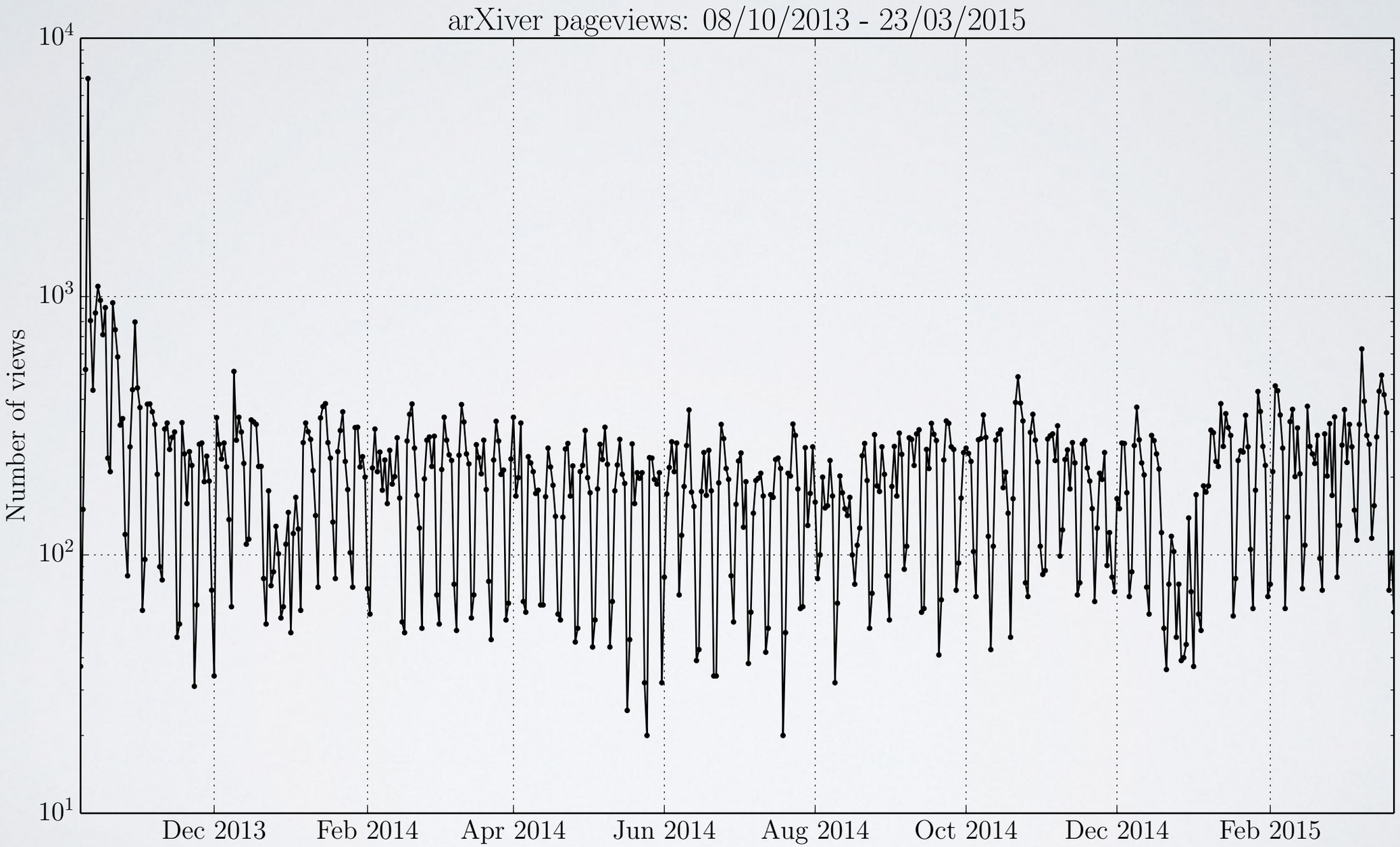
Current usage of arXiver

United States	29,680	Chile	2,159
Australia	11,786	South Africa	2,079
Netherlands	11,074	South Korea	1,365
Germany	10,006	Austria	1,343
United Kingdom	8,541	Switzerland	1,303
Japan	7,502	India	1,247
Spain	5,533	Czech Republic	1,209
Canada	4,208	Brazil	1,118
France	3,986	Sweden	1,057
Colombia	3,421	Russia	1,001
Italy	3,270	...	

= 124857 total views

Current usage of arXiver

Current usage of arXiver



The role of the arXiver user

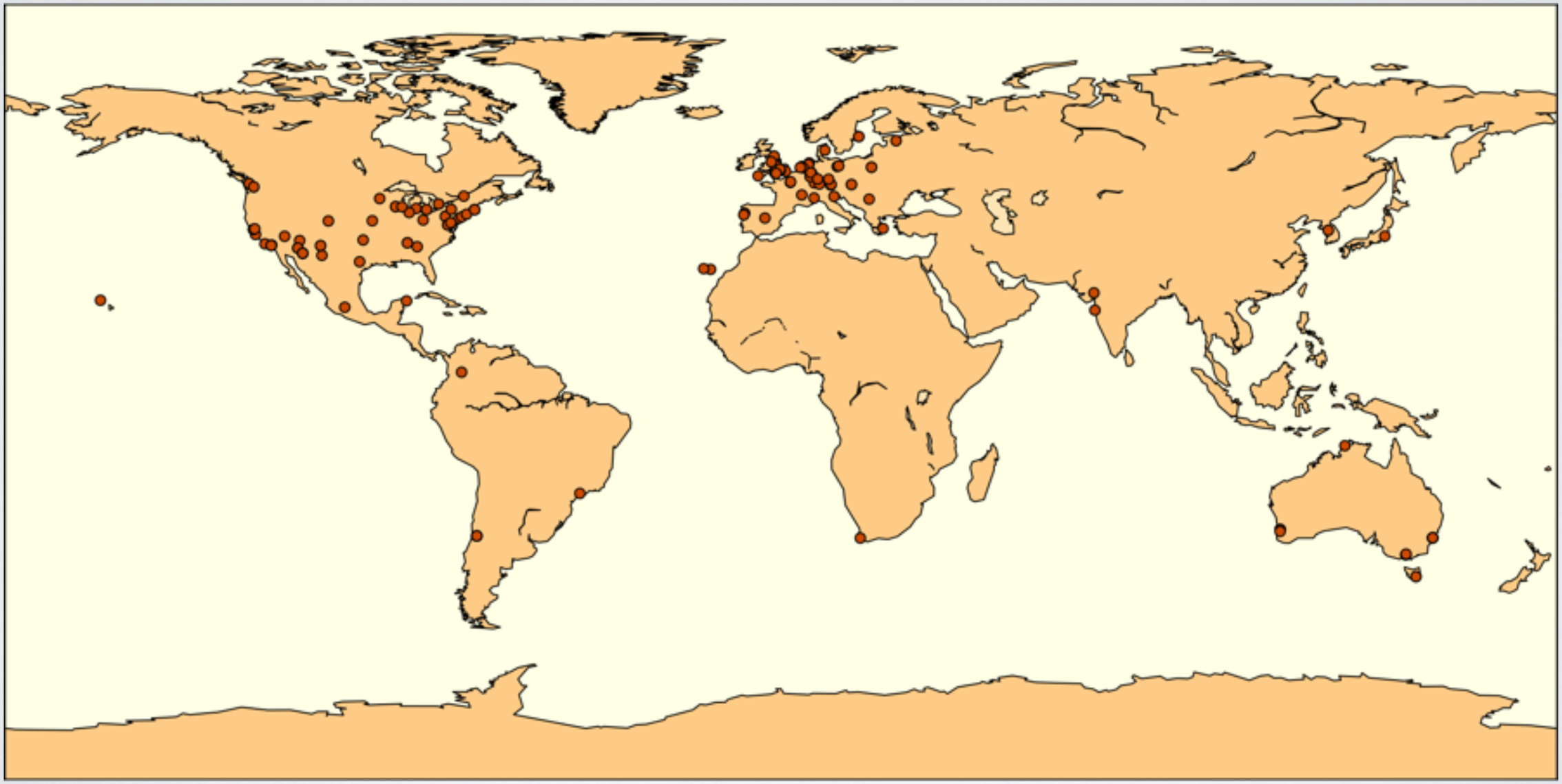
- We see arXiver as a tool for astronomers

- In order to gauge the current user experience of arXiver, we released an arXiver User Survey

- This survey had three goals:
  1) To determine opinion on showing author names in posts
  2) To identify the optimal path for future development
  3) To assess the response of astronomers to arXiver

- Overall we had 158 unique responses from around the world!

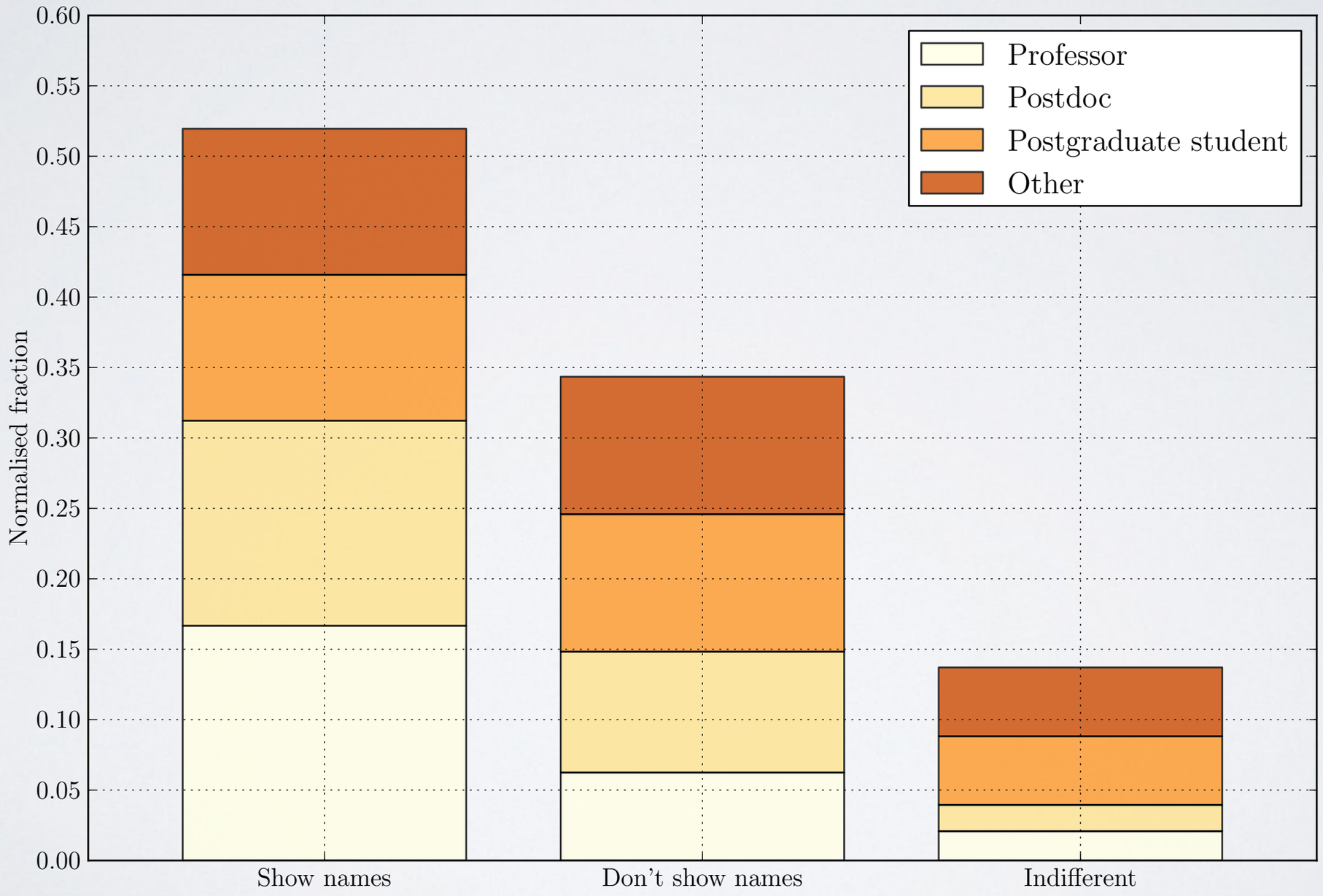
The arXiver user survey



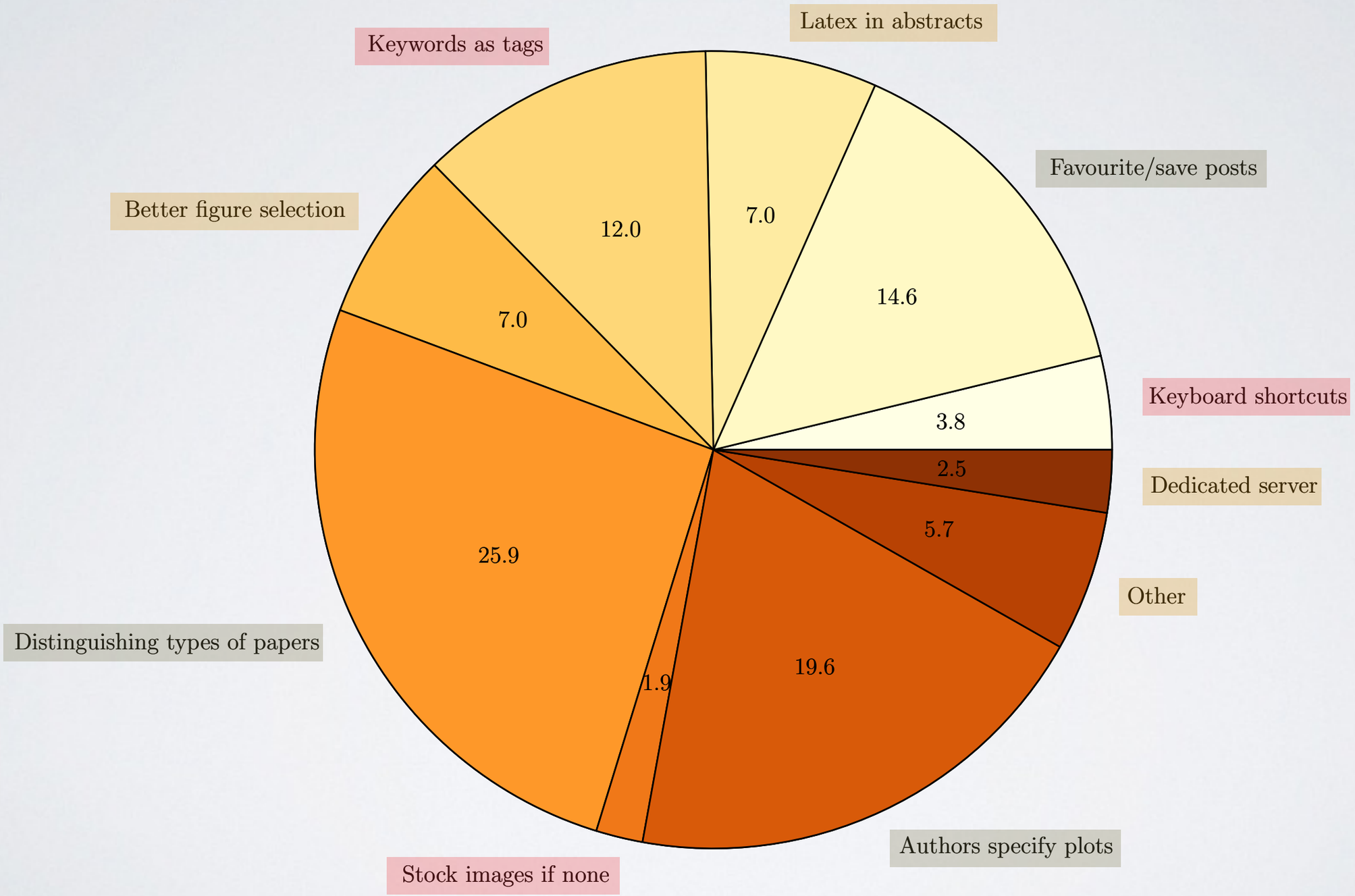
The question of author names

- At the start, we chose not to include author names in posts
- This was done with the goal that arXiver readers could judge each post based on its scientific content, not knowledge of authors (avoiding any unconscious or conscious bias)
- There are convincing arguments on both sides
  - avoiding bias, focusing on science
  - giving credit/exposure, familiarity with authors
- Part of the arXiver User Survey examined this question

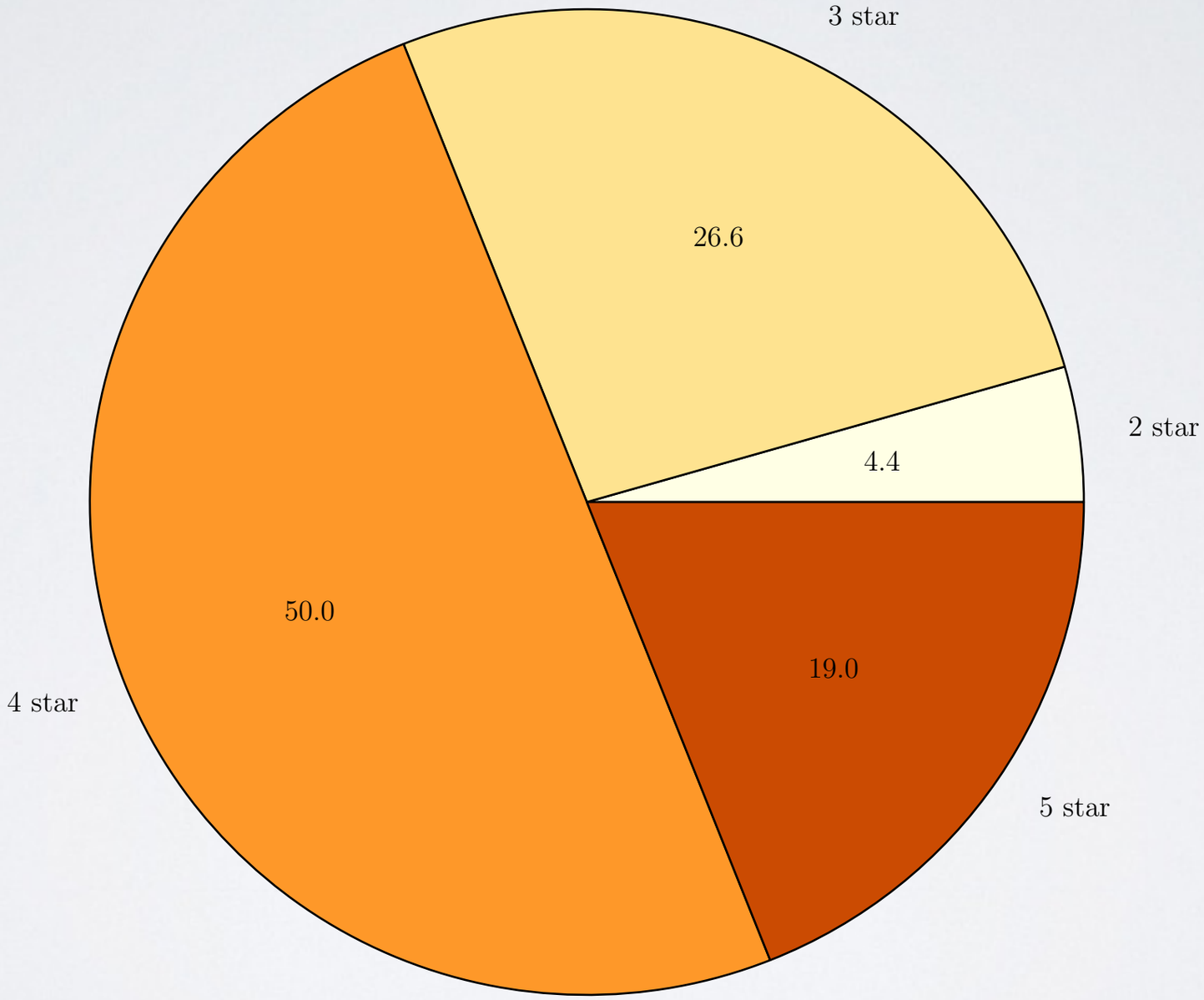
The question of author names



Paths for future developments



The 2013 experience of arXiver



The future of arXiver

- arXiver is at 207% storage capacity (6.2 GB) - need to move to a server!
- The arXiver User Survey results helped us determine what to focus on developing next, as well as answering the author question
- There may be potential for future collaboration with other open access scientific data projects in astronomy - arXiv? Other sciences?
- Fourier spectral occupancy is an effective way to determine the complexity of an image, and may be applicable beyond arXiver
- Python, combined with existing libraries, can be used to provide useful new services to aid astronomers in their work

Come talk to me!

- If you have ideas for new features or the future of arXiver, and how we can make it more useful for astronomers
- If you want to know more about any of the packages I mentioned, or seek advice on similar Python-web stuff
- If you know about transitioning from wordpress.com to wordpress.org and how to make it go smoothly
- If you are interested in Fourier spectral occupancy*

*see also: http://tftp.wikidot.com/

http://arxiver.net

Thanks for listening!