2018 Annual Report

Astrophysics Source Code Library

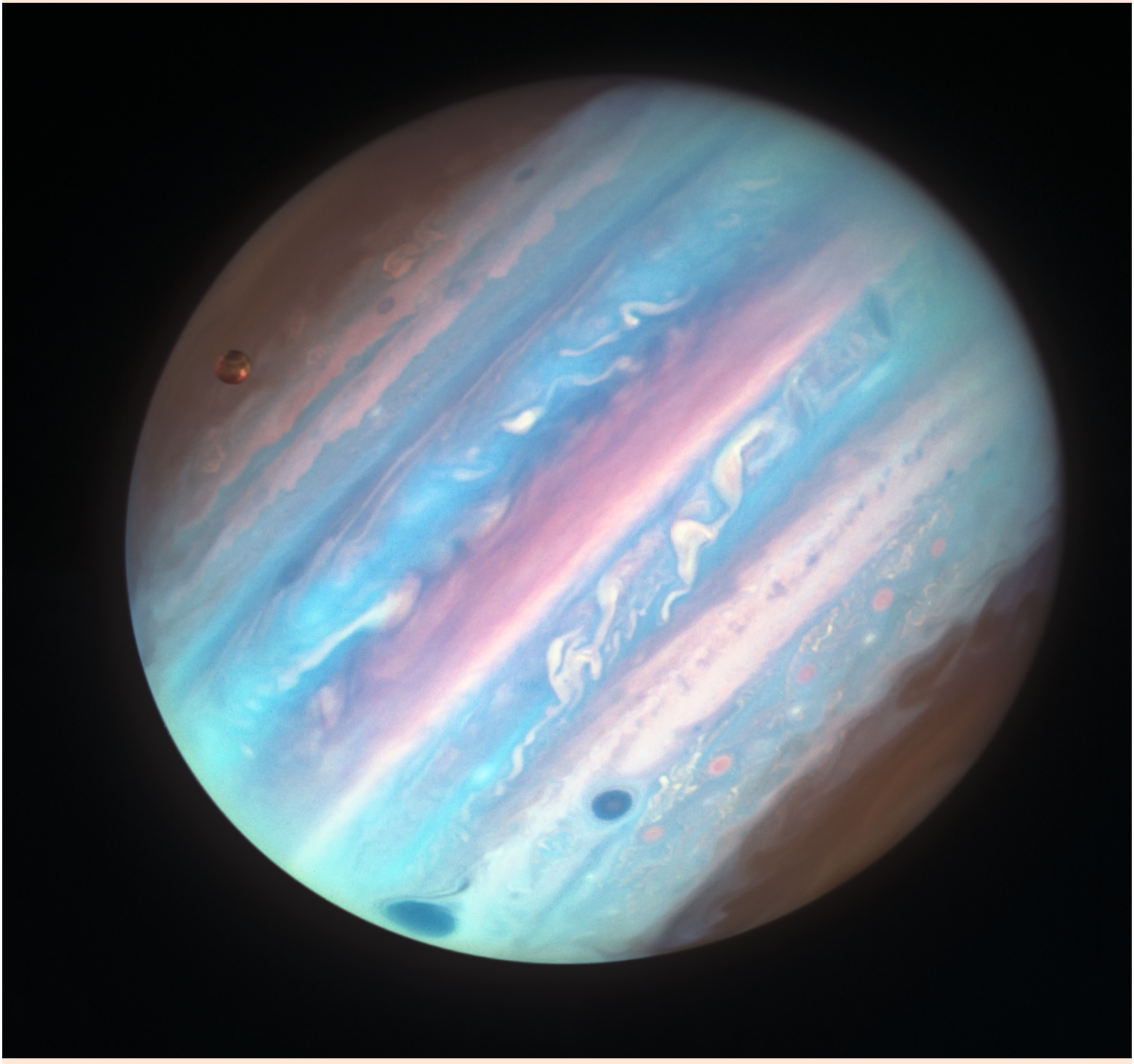


Table of Contents

Written by Alice Allen, ASCL Editor; February 2019
The cover image is Jupiter and Ganymede in Near-UV and Visible Blue Light by Judy Schmidt; used with permission.
Instrument: Hubble Space Telescope
This image represents Jupiter and Ganymede as they would have appeared on 2017-02-02 at 12:16 UTC.
https://www.flickr.com/photos/geckzilla/39448899094/in/dateposted/

Background

The Astrophysics Source Code Library (ASCL), founded in 1999 by Robert Nemiroff (Michigan Technological University) and John Wallin (Middle Tennessee State University), is a free online registry and repository for source codes of interest to astronomers and astrophysicists. All ASCL source codes have been used to generate results published in or submitted to a refereed journal. The ASCL website (http://ascl.net) is housed at Michigan Technological University; the site offers entries describing and linking to over 1800 source codes used in astronomy research, information about the resource itself and generally about research software, a news blog, and a discussion forum. The site links in our entries are actively curated and our link checkers consistently show that greater than 96.5% of our links are healthy. Even this high percentage is artificially low, as some links do not respond well to link checking yet are working.

2018 Growth

The number of codes indexed in ASCL grew an average of 21 entries per month, up from the average growth of 17.3 codes per month over the 2015-2017 timeframe and the 15.8 codes per month average for 2017. 251 codes were added in 2018. At the end of December, the ASCL had 1851 code entries.

Code submissions have remained healthy; 92 of the 251 codes (37%) assigned IDs in 2018 were submitted. This is a 31% increase over the 70 codes that were submitted and accepted in 2017. Fourteen codes submitted did not meet the criteria for inclusion in the ASCL; it is likely a subset of these will eventually meet our criteria and be assigned ASCL IDs.

Citations to the ASCL have also increased, up 59% in 2018 over 2017. Citations are discussed more fully in the section "***Impact on the community***."

New capabilities and features

In February 2018, the ASCL was added to the identifiers.org registry; the ASCL's record is available at https://identifiers.org/ascl. Also in February, the ASCL administration dashboard was made more robust with addition of statistics for the number of codes missing "Preferred citation" information and the number of codes missing information in both the "Appears in" and "Described in" fields. Having this information available on this dashboard allowed us a quick way to check our progress on two 2018 initiatives: increasing the number of entries with "Preferred citation" information, and continuing our work on disambiguating article links. Between our last annual report and this one, the number of entries with "Preferred citation" information has increased 392%, from 130 entries to 639. This initiative will continue in 2019. Disambiguating article links was a higher priority initiative for us in 2018 than was adding "Preferred citation" information, and we made great progress on it, having only 235 entries left with ambiguous article links by the end of the year and completion expected by the end of the first quarter in 2019.

We added a wiki to the ASCL site in April; this is currently used for tracking our progress on various projects and random shared notes between editors and developers. We are open to using the wiki for public information, but do not at this time have any idea as what to share there that is not available elsewhere on the site.

ADS held a Hack Day in Boston in May 2018; by the time of that event, over 50% of the ASCL's article links had been disambiguated into "described in" and "used in" article links. One of our projects for Hack Day was to work with ADS staff to allow them to ingest disambiguated links via a file created for bulk upload of this information; this project was successfully completed at the Hack Day, and these new links were displayed on ADS for the first time the following day. In June, our data feed report to ADS was changed to provide disambiguated link information along with new and changed record information.

Title: spec2d: DEEP2 DEIMOS Spectral Pipeline

Authors: Cooper, Michael C. ; Newman, Jeffrey A. ; Davis, Marc ; Finkbeiner, Douglas P. ; Gerke, Brian F.

Publication: Astrophysics Source Code Library, record ascl:1203.003

Publication Date: 03/2012

Origin: ASCL

Keywords: Software

Bibliographic Code: 2012ascl.soft03003C

Abstract

The DEEP2 DEIMOS Data Reduction Pipeline ("spec2d") is an IDL-based, automated software package designed to reduce Keck/DEIMOS multi-slit spectroscopic observations, collected as part of the DEEP2 Galaxy Redshift Survey. The pipeline is best suited for handling data taken with the 1200 line/mm grating tilted towards the red (lambda_c 7800Å). The spec2d reduction package takes the raw DEIMOS data as its input and produces a variety of outputs including 2-d slit spectra and 1-d object spectra.

Associated Articles

Source Software Paper 1

Title: spec2d: DEEP2 DEIMOS Spectral Pipeline

Authors: Cooper, Michael C. ; Newman, Jeffrey A. ; Davis, Marc ; Finkbeiner, Douglas P. ; Gerke, Brian F.

Publication: Astrophysics Source Code Library, record ascl:1203.003

Publication Date: 03/2012

Origin: ASCL

Keywords: Software

Bibliographic Code: 2012ascl.soft03003C

Abstract

The DEEP2 DEIMOS Data Reduction Pipeline ("spec2d") is an IDL-based, automated software package designed to reduce Keck/DEIMOS multi-slit spectroscopic observations, collected as part of the DEEP2 Galaxy Redshift Survey. The pipeline is best suited for handling data taken with the 1200 line/mm grating tilted towards the red (lambda_c 7800Å). The spec2d reduction package takes the raw DEIMOS data as its input and produces a variety of outputs including 2-d slit spectra and 1-d object spectra.

Associated Articles

Source Software Described in

Article-software cross-matching in ADS
Top: ambiguous link; bottom: disambiguated link

A new field, "Keywords," was added to the ASCL database in July as part of our funded ADAP project; the keyword "NASA" is being added to ASCL records for NASA software. ADS first ingested keyword-tagged records in October, making it possible for the first time to search ADS for doctype:software keyword:NASA and get results. More information about this project is provided in the ***Support and funding*** section.

Research

In 2017, ASCL Advisory Committee Chair Peter Teuben, developer P.W. Ryan, and Editor Alice Allen conducted a preliminary study on software availability and link persistence in astrophysics, which was submitted to and accepted by *The Astrophysics Journal Supplement*. This paper was published in May 2018 in a special issue on data. Ryan conducted additional research on link persistence in astrophysics and presented it at the Astronomical Data Analysis Software and Systems (ADASS) conference held in College Park, MD, in November 2018. The poster for this presentation can be found at the end of this report.

Support and funding

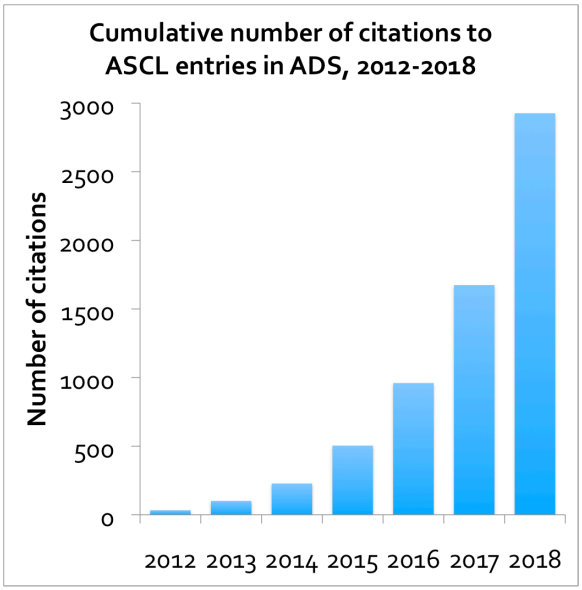
The Heidelberg Institute in Theoretical Studies (HITS) provided €6,000 in funding at the end of 2017 for work in 2018; this support enabled the ASCL's presence at various conferences and participation in meetings on research software such as the joint Royal Astronomical Society National Astronomy Meeting and the European Week of Astronomy and Space Science (the European Astronomical Society's large annual meeting) in Liverpool and the International Astronomical Union General Assembly in Austria, and outreach to organizations, including ASTRON, NRAO, and ESO. The ASCL is deeply grateful for this support.

Michigan Technological University (Houghton, MI) hosts the ASCL website and provides technical support, and the University of Maryland (College Park, MD) provides office space and other university services including library access; the ASCL assigns DOIs through an arrangement with the University of Maryland Libraries. The ASCL is very grateful for this valuable ongoing support.

In May 2017, Peter Teuben and Alice Allen submitted a proposal for NASA funding under its ADAP solicitation, this to improve the visibility and discoverability of NASA software through the ASCL. In late January 2018, we were delighted to learn that our proposal was fully funded at $162K for two years. Work started on the project in July 2018, and our first results were available in ADS in October 2018.

Impact on the community

The ASCL uses various measures to try to determine its impact on and use by the community. Other than wanting software to be cited in a trackable way, the ASCL is agnostic on how codes are cited. That said, we use citation numbers as a metric to learn how and whether the ASCL is being used. Citations continue to increase at a faster rate than growth in the number of code entries. By the end of 2018, 595 ASCL entries collectively had 2926 citations, meaning 32% of the 1851 codes indexed in ADS at that time had citations, up from 28% in December 2017; the number of citations grew by 59%, and the number of entries cited grew 35%, from 442 to 595.



Increasingly, authors submit their software to the ASCL; 106 code entries were submitted in 2018, with 92 of these accepted into the ASCL. This is a 31% increase in the number of submitted codes over 2017, which saw 85 submitted code entries. We take this as an indication that software authors see value in registering their codes with the ASCL. Interestingly, authors are not the only submitters of software to the ASCL; occasionally, a data editor or a user has submitted a code to enable citation of the software.

ASCL records continue to be indexed quarterly by the Web of Science's Data Citation Index; unfortunately, as of this writing, Clarivate has reported an internal problem with ASCL entries and we are unable to pull citation information from that site at this time. Clarivate is working on the problem and we will issue an update to this report when it is resolved.

People

Alice Allen and Kimberly DuPrie (Space Telescope Science Institute) are Editor and Associate Editor, respectively. Judy Schmidt provides development and design work for the ASCL, and P.W. Ryan provides development work and support for citation and other statistics tracking.

The Advisory Committee members in 2018 were:

Peter Teuben, University of Maryland, *Chair*
Bruce Berriman, Infrared Processing and Analysis Center/Caltech
Jessica Mink, Center for Astrophysics
Robert Nemiroff, Michigan Technological University
Lior Shamir, Lawrence Technological University
Keith Shortridge, Knave and Varlet, AU
Mark Taylor, University of Bristol, UK
John Wallin, Middle Tennessee State University
Rein Warmels, European Southern Observatory, DE

2018 Highlights

January	Held a Special Session on Astronomy Software Publishing at the 231st AAS meeting. Session report: http://ascl.net/wordpress/2018/01/17/report-on-the-astronomy-software-publishing-special-session-at-aas231/ Presented the ASCL in a poster: http://ascl.net/wordpress/2018/01/10/ascl-poster-at-aas-231/ Pre-print of research paper appears on arXiv: https://arxiv.org/abs/1801.02094
February	Lior Shamir, Bruce Berriman, Peter Teuben, Robert Nemiroff, and Alice Allen submitted the white paper "Best Practices for a Future Open Code Policy: Experiences and Vision of the Astrophysics Source Code Library" on the ASCL's behalf to the National Academy of Sciences, Engineering, and Medicine's Best Practices for a Future Open Code Policy for NASA Space Science Project Committee Made the admin dashboard more robust
April	EWASS/NAM Software in Astronomy Symposium held, including software publishing special interest group meeting SPSIG blog post | SPSIG Google doc Wiki is added to ASCL Organizational outreach talks presented at the University of Amsterdam and ASTRON
May	ApJS issue with our "Schroedinger's Code" research article published Activities at ADS Hack Day include ingestion of disambiguated ref links by ADS and new display of this information under "Associated Articles"
June	Organizational outreach talk presented at NRAO/UVA Report for ADS changed to report disambiguated ref links
July	Started work on NASA ADAP project

August	2015 (Sydney) ADASS proceedings published; this includes two ASCL-related papers: https://ui.adsabs.harvard.edu/#abs/2018ASPC..512..517A https://ui.adsabs.harvard.edu/#abs/2018ASPC..512..517A Presented poster at IAU General Assembly
September	Submitted proposal for next year's EWASS Presentation on the ASCL at AstroInformatics conference (by invitation; Rein Warmels presented) Organizational outreach talk presented at Lyon Observatory Informal outreach discussion at ESO (Garching)
October	First ingestion of NASA-tagged ASCL entries into ADS Presentation on ASCL and software citation at Deep Learning for Multimessenger Astrophysics: Real-time Discovery at Scale conference at UIllinois, NCSA (by Bruce Berriman) Organizational outreach talk presented at University of Delaware Participated in AAHEP meeting at arXiv at Cornell
November	Participated in ADASS; presented oral talk "Receiving Credit for Research Software" and coordinated unconference BoF session: "I want to talk about..."; presented research poster on "Schroedinger's Code" paper

2018 Plans Revisited

- Create ASCL index in 2017 ADASS proceedings
  - Proceedings not yet available; task deferred until document is available
- Make the administrator dashboard more robust
  - Completed; changes included addition of number of entries missing "Preferred citation", and "Appears in" and "Described in" information
- Present ASCL at at least two conferences
  - Completed; ASCL was presented at AAS in January, NAM/EWASS in April, International Astronomical Union General Assemby in August, AstroInformatics in September, Deep Learning for Multimessenger Astrophysics in October, and ADASS in November
- Present ASCL at at least six institutions
  - Completed; ASCL was presented at University of Amsterdam, ASTRON, ESO, Lyon Observatory, NRAO, ESO, and University of Delaware
- Sustain reasonable growth in number of entries (200-220 additions)
  - Completed; 251 codes were added in 2018
- Implement a mirror site
  - Still outstanding; not completed
- Add another two members to the Advisory Committee
  - Still outstanding; not completed

2019 Goals

- Create ASCL index in 2017 and 2018 ADASS proceedings
- Complete the disambiguation of article links and modify screens and reports accordingly
- Present ASCL at at least two conferences
- Present ASCL at at least six institutions
- Sustain reasonable growth in number of entries (200-220 additions)
- Offer auto-generated codemeta.json and/or citation file format files to software authors
- Complete outstanding plans from previous years
  - Add two members to the Advisory Committee
  - Implement a mirror site

Press, posters, and papers

The Astrophysics Source Code Library by the numbers
Alice Allen *et al*, American Astronomical Society, AAS Meeting #231, January 2018, id.150.10
https://adsabs.harvard.edu/abs/2018AAS...23115010A

Schroedinger's code: Source code availability and transparency in astrophysics
PW Ryan, Alice Allen, and Peter Teuben, American Astronomical Society, AAS Meeting #231, January 2018, id.150.28
https://adsabs.harvard.edu/abs/2018AAS...23115028R

Irreproducible astronomy
Physics Today, April 4, 2018
https://physicstoday.scitation.org/do/10.1063/PT.6.1.20180404a/full/

GPUs Mine Astronomical Datasets For Golden Insight Nuggets
The Next Platform, April 16, 2018
https://www.nextplatform.com/2018/04/16/gpus-mine-astronomical-datasets-for-golden-insight-nuggets/

Schroedinger's code: A preliminary study on research source code availability and link persistence in astrophysics
Alice Allen, Peter J. Teuben, P. Wesley Ryan, Astrophysical Journal Supplement Series, Volume 236, Number 1, 2018 May
http://iopscience.iop.org/article/10.3847/1538-4365/aab764/meta

Best Practices for a Future Open Code Policy: Experiences and Vision of the Astrophysics Source Code Library
Lior Shamir, Bruce Berriman, Peter Teuben, Robert Nemiroff, Alice Allen
https://arxiv.org/abs/1802.00552
White paper submitted to the National Academies of Sciences, Engineering, and Medicine's Best Practices for a Future Open Code Policy for NASA Space Science Project Committee
http://sites.nationalacademies.org/SSB/CurrentProjects/SSB_178892

Computational astrophysics for the future
Science, September 7, 2018
http://science.sciencemag.org/content/361/6406/979

AAS Poster, January 2018

The Astrophysics Source Code Library by the numbers

Alice Allen, Peter Teuben, G. B. Berriman, Kimberly DuPrie, Jessica Mink, Robert Nemiroff, PW Ryan, Judy Schmidt, Lior Shamir, Keith Shortridge, John Wallin, Rein Warmels

Founded in 1999, the Astrophysics Source Code Library (ASCL) is a free registry of source codes used in research that is or has been submitted for peer review.

The ASCL is indexed by the SAO/NASA Astrophysics Data System (ADS) and Web of Science.

ASCL entries can be cited by the unique ascl ID assigned to each code.

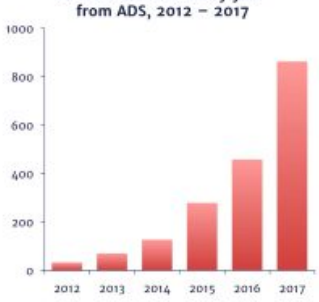
Number of codes: 1600

Percentage of entries submitted by their authors since mid-2014: 20%

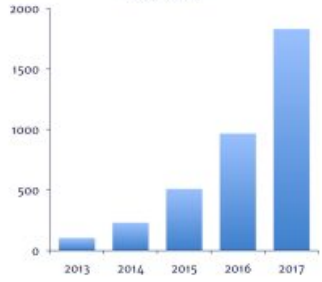
Percentage increase in the number of submissions by authors from 2016 to 2017: 23%

Number of times published code entries have been viewed since July, 2014: 1.8M

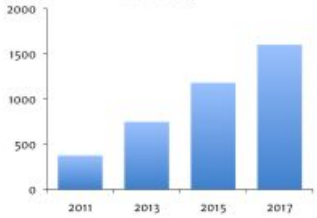
Number of citations by year from ADS, 2012 – 2017



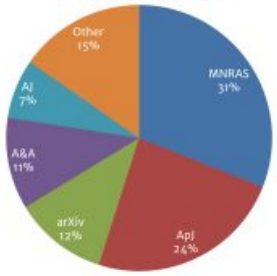
Cumulative citations by year, 2013-2017



Cumulative number of codes, 2011-2017

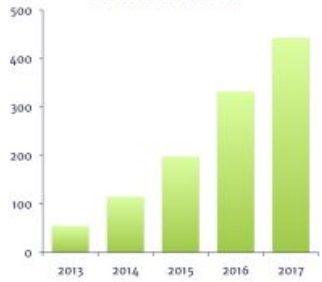


Citations by journal as of January 4, 2018



*Number of hours spent making posters about the ASCL: 5,000,060**

Cumulative number of code entries with citations



Number of code entry views, highest 10

McScatter: Three-Body Scattering with Stellar Evolution	8608
COSMICS: Cosmological initial conditions and microwave anisotropy codes	7660
Turbospectrum: Code for spectral synthesis	6982
BHSKY: Visual distortions near a black hole	6420
NEMO: A Stellar Dynamics Toolbox	5493
ACORNS-ADI: Algorithms for Calibration, Optimized Registration and Nulling the Star in Angular Differential Imaging	4992
PyMGC3: Finding stellar streams in the Galactic Halo using a family of Great Circle Cell counts methods	4541
Transit Analysis Package (TAP and autoKep): IDL Graphical User Interfaces for Extrasolar Planet Transit Photometry	4362
Starlink: Multi-purpose Astronomy Software	4349
IM3SHAPE: Maximum likelihood galaxy shear measurement code for cosmic gravitational lensing	4009

Journals with citations to ASCL entries (partial list)

arXiv (arXiv)	Journal of Physics: Conference Series (JPhCS)
Astronomical Journal (AJ)	Monthly Notices of the Royal Astronomical Society (MNRAS)
Astronomical Society of the Pacific Conference Series (ASPC)	Nature (Natur)
Astronomy & Astrophysics (A&A)	Nature Astronomy (NatAs)
Astronomy & Computing (A&C)	PASA (PASA)
Astrophysical Journal (ApJ)	PhD Thesis (PhDT)
Astrophysical Journal Supplement (ApJS)	Physical Review C (PhRvC)
Classical and Quantum Gravity (CQGra)	Physical Review D (PhRvD)
Galaxies (Galax)	Physical Review Letters (PhRvL)
IAU Symposium (IAUS)	Publications of the Astronomical Society of Japan (PASJ)
Information Bulletin on Variable Stars (IBVS)	Publications of the Astronomical Society of the Pacific (PASP)
Journal of Cosmology and Astroparticle Physics (JCAP)	Science (Sci)

ACKNOWLEDGEMENTS

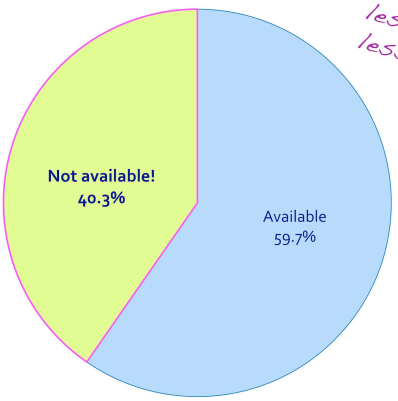
We thank the Heidelberg Institute for Theoretical Studies for financial support, and Michigan Tech and the University of Maryland for general support.

Heidelberg Institute for Theoretical Studies | HITS | Michigan Tech

* Probably not really, it just feels that way

Improve your research reproducibility with the Astrophysics Source Code Library (ascl.net)

We sampled 166 research papers; over **40%** of the software used in the research was **not available** as source code!



Availability of unique codes

This makes research less transparent and less reproducible!

Why release source code?

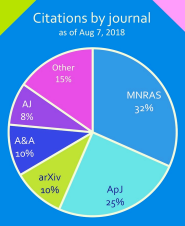
- Uphold research transparency
- Increase reproducibility of your research
- Show confidence in your results
- Improve efficiency in the discipline

Researcher benefits

- Linking of research and software on ASCL and ADS
- Preferred citation information for many entries
- Curated site/repos links for useful downloadable software

What is the ASCL?

- Free online citable registry of source codes used in astro research
- Indexed by Astrophysics Data System (ADS) & Web of Science (WoS)
- Actively curated
- Nearly 1800 codes registered



NASA-funded project goals

- Add keyword field to ASCL
- Tag NASA codes with keyword
- Pass keywords to ADS
- Allow keyword filtering on ASCL

Benefits of registering software with the ASCL

- Code is citable on its own merit in accordance with Force11 Software Citation Guidelines and, increasingly, journal policies
- Software is more discoverable; entries also appear in ADS & WoS
- Citations by ASCL ID are tracked by ADS, unlike repository DOIs

The ASCL is supported by
Heidelberg Institute for Theoretical Studies
Michigan Tech University
University of Maryland
NASA

Alice Allen (ASCL/University of Maryland), Robert J. Nemiroff (Michigan Technological University), Peter J. Teuben (University of Maryland)

ADASS Poster, November 2018

Schrödinger's code: Research transparency in code and link accessibility

P. Wesley Ryan (ASCL, wes@ascl.net), Alice Allen (ASCL/UMD), Peter Teuben (UMD)

Why do we care whether source code is available and hyperlinks in papers work?

Astrophysics, like most disciplines, depends on software.

The increase in reliance on software has led to a decrease in research transparency and reproducibility, as computational methods that underlie and enable research are frequently not made available.

Astrophysics papers frequently contain hyperlinks to external resources. These hyperlinks may refer to code, data, or other information crucial to the paper.

Failure of these links over time may decrease research transparency, understanding, and reproducibility.

What we did

We sampled the papers published in Astronomy & Astrophysics (A&A) in 2015 to determine what percentage of astrophysics research code is available.

We extracted embedded hyperlinks from all papers published in A&A and the Astrophysical Journal (ApJ) in 2015 to measure the persistence of these links.

How we did it

For source code

We examined 10% of research articles published in *A&A* in 2015 for definite or possible use of computational methods and built a file of this information.

418 of our records included readily searchable information about software, such as a name, a URL, or an ASCL entry.

We searched for download sites for the 285 unique codes in the 418 records and assigned each code to a category corresponding to its accessibility.

For hyperlinks

We extracted the clickable hyperlinks from 1,669 papers published in A&A into a database, filtering out hyperlinks to email addresses, DOIs, and nine frequently-cited Web sites that we knew to be available, to build a file of 2,558 hyperlinks.

We tested the extracted hyperlinks four times on four different dates in September and October 2017 from different locations and computers.

One year later, we tested the same set of links again in order to measure link persistence over time. We also extracted, filtered, and tested links in the 3,050 papers published in ApJ in 2015.

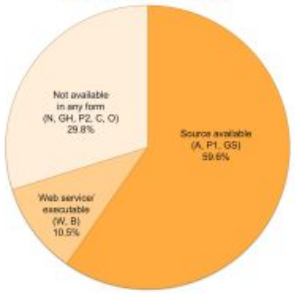
Our suite of scripts, LExTeS, is publicly available at https://ascl.net/1711.018.

What we found out

41% of the software used did not offer source code. Wow!!

30% of software used was not available in any form.

Availability of unique codes



285 unique codes were used in our sample papers.

80% of codes were used in only one paper in our sample; nearly all of these were used in previous or subsequent research.

58% offer downloadable source code (categories A and GS)
2% came from *Numerical Recipes* (category P1)
11% were available as an executable or online service (categories B and W)

Software categories

Category	Description
A	source code readily available
P1	source code available for purchase (Numerical Recipes)
GS	source code soft-gated
W	web service
B	executable file (binary or compiled code)
N	not found
GH	source code hard-gated
P2	commercial software (such as Mathematica or IDL)
C	source code is available only to collaborators
O	other

Unique code category summary

Category	Number of unique codes	Percentage
A	162	56.8%
P1	6	2.1%
GS	2	0.7%
W	16	5.6%
B	14	4.9%
N	63	22.1%
GH/P2/C/O	22	7.7%

Link accessibility summary (excluding FTP)

Links (time of testing)	Up	Down	Inconsistent
A&A (Oct. 2017)	86.9%	10.6%	2.5%
A&A (Oct. 2018)	86.1%	12.7%	1.2%
ApJ (Oct. 2018)	83.6%	14.6%	1.8%

In 2017, 9.5% to 13.2% of hyperlinks tested were inaccessible, 10.6% of hyperlinks were consistently inaccessible, and 2.5% of hyperlinks were inconsistently accessible in the period between the first and last runs of our script.

Only 0.8% of links that consistently worked in 2017 were inaccessible in 2018, but 1.3% of previously inconsistent links became inaccessible, resulting in a total of 12.7% broken links.

Conclusions

The use of open source packages in multiple papers weights the overall "availability index" of source code favorably upward and has a positive impact on the transparency of astrophysics research.

Link inaccessibility does not happen at a constant rate.

Future directions and similar research in other disciplines

We plan to examine articles in other journals and published in other years to see whether availability of source code changed over time, and test our links dataset periodically to determine link persistence over a longer time period.

Collberg & Proebsting (2016) examined source code availability in computer systems research and found that only 56% of software-backed articles had available source code, and Howison & Bullard (2016) found that "software is frequently inaccessible" and "between 24%-40%" of biology articles provide source code.

Mangul *et al.* (2018) found that 26% of computational biology software tools were not accessible through the URLs referenced in published papers, and 57% of the tools sampled for a usability test could not be installed by strictly following the instructions provided in the manual.

Mangul *et al.* also found that 32% of software linked in computational biology papers published before 2012 was inaccessible by the provided URL, whereas only 14% of software linked in papers published after 2012 was.

Citations
Allen A., Teuben, P. J. & Ryan, P. W. 2018 ApJS, 236, 10
Collberg, C., & Proebsting, T. A. 2016, Commun. ACM, 59, 62. http://doi.acm.org/10.1145/2812803
Howison, J., & Bullard, J. 2016, Journal of the Association for Information Science and Technology, 67, 2137. http://dx.doi.org/10.1002/asi.23538
Mangul, S., et al. 2018, bioRxiv pre-print. http://dx.doi.org/10.1101/452532
Ryan, P.W., 2017, LExTeS, Astrophysics Source Code Library, ascl:1711.018