Astro 2020 ‘Infrastructure Activity’ White Paper: Citizen Science as a Core Component of Research Infrastructure

Date: July 6, 2019

Principal Author: Laura Trouille (Zooniverse; The Adler Planetarium, Northwestern University trouille@zooniverse.org, 312-322-0820)

Co-authors: Chris Lintott (Zooniverse; University of Oxford), Lucy Fortson (Zooniverse; The University of Minnesota - Twin Cities)

Abstract

Since the 2010 Astronomy and Astrophysics Decadal Survey, enormous progress has been made in establishing citizen science (or the public participation in scientific research) as a reliable and effective tool for processing increasingly large astronomical datasets. The Zooniverse model of an open platform, available to researchers for free, has been instrumental in enabling projects ranging from TESS to LSST, and ZTF to LIGO to apply a citizen science approach to data analysis. In integrating machine learning into the citizen science system, we have learned that humans and machines working in combination addresses the classification problem with greater facility than machines alone, despite recent increases in machine learning capability. The added ability for serendipitous discovery, which platforms like the Zooniverse and their associated discussion forums provide, further ensures that even for very large datasets, citizen science will retain its place as a key element of astronomical data analysis throughout the decade to come. To maximize the return provided by such projects, we recommend the following actions be taken: 1) support and encourage the use of shared infrastructure for online citizen science; 2) there should be a default move to make tools and, where plausible, data, developed by missions and large projects available to all; 3) support for ongoing maintenance of citizen science platforms as infrastructure; and 4) continued incorporation of explicit language in funding solicitations that encourages the use of citizen science.

1 Introduction

The involvement of the general public in astronomical research has a history which dates back to Edmund Halley's study of the 1715 total solar eclipse, which included observations from observers scattered across the British Isles [23]. Organizations such as the American Association of Variable Star Observers (AAVSO) have, for more than a century, coordinated and collated the efforts of amateur observers, who have made significant contributions to the study of our Solar System, to that of variable stars and time domain astronomy more generally. The widespread adoption of digital technologies, which accompanied the general adoption of surveys producing large, open datasets by the professional community, has driven a great expansion in the capacity of such 'citizen science' projects over the last 15 years. These new projects asked large crowds of volunteers to contribute annotation or analysis, rather than observation. Since the 2010 'New Worlds, New Horizons in Astronomy and Astrophysics' decadal survey recognized citizen science 'as a new and promising method of public engagement', this way of conducting science has become a standard tool in astronomers' arsenal, adopted by a wide range of missions and surveys [21].

This expansion has driven and in turn been influenced by a much more general increase in interest in the capacities of citizen science for both research and engagement; the America COMPETES Reauthorization Act of 2010, the Crowdsourcing and Citizen Science Act of 2016 and the 2019 White House report by the Office of Science and Technology Policy [22] highlight the critical role that citizen science can play in facilitating projects that would otherwise be beyond the reach of professionals. The largest set of projects involving data analysis is hosted by the Zooniverse, which supports projects in disciplines including astronomy, biomedical science, climate science, ecology and the humanities.

Zooniverse, which offers a reliable, flexible and scalable API and an online project builder for scientists considering citizen science, grew from a single project, Galaxy Zoo, launched in 2007 [13]. Projects as different as TESS and LSST, ZTF and STEREO, have depended on Zooniverse to support their citizen science efforts. Indeed, Zooniverse projects have helped with analysis of data across the full electromagnetic spectrum - from radio [1] to gamma rays [2] - as well as with multi-messenger astrophysics (e.g., LIGO's GravitySpy.org [11]) and cosmological simulations [4]. In a companion submission on 'EPO Vision, Needs and Opportunities' we provide recommendations for Citizen Science for EPO, looking broadly at the Citizen Science landscape. Here, we use the specific example of Zooniverse to argue firstly that citizen science will be needed in the scientific landscape of the next decade, and then that funding agencies should support a common platform for use by projects of all scales.

2 Classification and the role of machine learning

Galaxy Zoo's catalogues of morphology have been used extensively by researchers over the last decade (see Watson & Floridi [30] for an analysis of this use) and have been foundational to progress in understanding galaxy formation. The careful analysis of classifications from multiple users [19, 20] produced a catalogue which matched expert classification properties, and the project team was able to move on to detailed morphological classifications of both the Sloan Digital Sky Survey and large HST surveys [32, 31, 26].

As well as being useful in their own right, however, large, labelled datasets are necessary for the effective use of modern machine learning methods, particularly deep learning using convolutional neural networks (CNN), which have been shown to be effective in a variety of astronomical domains. It is therefore not surprising that Zooniverse projects have been used to generate such data; an early transient astronomy project ('Galaxy Zoo: Supernova') which used data from the Palomar Transient Factory was retired from the system after volunteers produced a large enough training set for machine learning to process the remaining data with confidence [27]. The Galaxy Zoo dataset was used to run a machine learning competition on the Kaggle platform, which provided an early demonstration of the astronomical utility of convolutional neural networks [5], and many other large datasets produced by the Zooniverse team have been used for similar purposes (e.g.[33]).

Given this context, it is tempting to argue that – despite the diversity of research projects currently relying on citizen science projects – continued rapid advances in machine learning coupled with the large datasets expected from future surveys will eventually close the parameter space of large data and human capability which has enabled the growth of citizen science in the last decade. However, a more recent set of projects demonstrates that this is not the case, and that there is great potential in combining human and machine classification [9, 14].

The current Supernova Hunters project provides an example of the power of this approach. Each week, thousands of images from the Pan-STARRS telescope are flagged by algorithms as potential supernovae, and those which are considered especially promising are reviewed by volunteers. Wright et al. [8] found that combining volunteer classifications with a trained CNN outperformed either approach operating in isolation; the machine provides a complete but typically impure sample, while human classifications provide a pure but incomplete sample which includes real examples that would otherwise be flagged as outliers. More recently, Walmsley et al. [29] built on previous, theoretical work by Beck et al. [10] to introduce to Galaxy Zoo a Bayesian neural network capable of predicting uncertainty in classification; this is being used to identify the galaxies whose classification by volunteers would lead to the greatest improvement in machine performance.

These hybrid methods, which allow training of machine learning to concentrate on the most common features in the dataset, but which preserve a role for citizen science, mean that we should expect citizen science to retain its place as part of astronomical data analysis throughout the decade to come. However, a focus on classification ignores some of the most important results to come from citizen science to date, which we discuss in the next section.

3 Serendipitous Discoveries

A significant advantage of volunteer review of large datasets is the ability to identify interesting and unusual objects. The ability of volunteers to go beyond their training is a key advantage over even well-trained automatic algorithms. For example, the 'Red ring' – a highly star forming lensed AGN [15, 16] – was discovered by volunteers on the SpaceWarps project who had been shown a training set consisting of blue lenses.

The kind of close scrutiny afforded by crowds of volunteers is also effective in finding the truly unexpected. Early examples from Galaxy Zoo include 'Hanny's Voorwerp' [7], a quasar-scale light echo which began an ongoing program of studying ionized gas around fading AGN ([17] and more), and the Green Pea galaxies, which are highly efficiently star forming low mass systems

[6]. In this latter case, citizen scientists took more of an active part in analyzing their find, going beyond simply identifying the unusual to conducting a systematic study [28].

This behavior is especially prevalent in astronomical projects with rich datasets and a mature ecosystem of tools which can be used by both professional and citizen scientists, such as the Planet Hunters project which drew on data from the Kepler and K2 (and now TESS) missions. For example, Planet Hunter volunteers were responsible for the identification of 'Boyajian's Star', KIC 8462852, now the prototype of a class of irregular dippers [12]. The most recent and extreme example of the class, EPIC 249706694, exhibits 28 transit-like dips during a 87 day period of observation, without any apparent periodicity [25], and was first identified by Planet Hunter volunteers. Several volunteers who started by participating in the main project now carry out research independently; the identification by Jacobs of a star with likely exocomet transits (reported in [24]) is an excellent example. Another example stems from the Backyard Worlds: Planet 9 project searching for brown dwarfs and other objects at the edge of our Solar System using NASA's Wide-field Infrared Survey Explorer (WISE) data. The project launched in Feb 2017 and to date has received over 5 million classifications from more than 150,000 participants. They have discovered over 100 new brown dwarfs so far, with an additional 1000 candidates awaiting confirmation (see [18, 3]). Approximately half of the candidates have been identified through the primary classification task within Zooniverse and the other half through WISEVIEW[1], a custom, add-on tool built by a Backyard Worlds volunteer, taking advantage of the open access to WISE data and metadata.

Serendipitous discovery is thus a key benefit of maintaining a citizen science approach to data analysis. We have argued in the previous section that humans and machines working in combination addresses the classification problem with greater facility than machines alone, despite recent increases in machine learning capability. The added ability to find the single most interesting object within a very large dataset, which platforms like the Zooniverse and their associated discussion forums and collaborations provide, is a further argument for ensuring that even the very large datasets expected in the next few years are inspected by volunteers.

4 Project growth

As rates and volumes of data steadily climb, the number of research teams wanting to build crowdsourcing projects has similarly grown. Over the past decade, Zooniverse has grown from a single, custom-built project in 2007 to a platform that has enabled over 150 projects (with 90 currently active projects) and engages over 1.7 million volunteers worldwide. The accelerated expansion of Zooniverse is a result of the launch in July 2015 of the Zooniverse Project Builder platform[2], which enables anyone to create an online citizen science project using a web browser-based toolkit.

The Project Builder supports common types of interaction including classification, multiple-choice questions, comparison tasks, text entry, marking and drawing tools, or any combination thereof. The Project Builder front-end is a series of forms and text boxes a researcher fills out to create the project's classification interface and website (Figure 2). All Project Builder projects come with a landing page, classification interface, discussion forum, and 'About' pages for content about the research, the research team, and results from the project.

[1]http://byw.tools/wiseview and https://youtu.be/23hXbfh1_aA
[2]https://www.zooniverse.org/lab

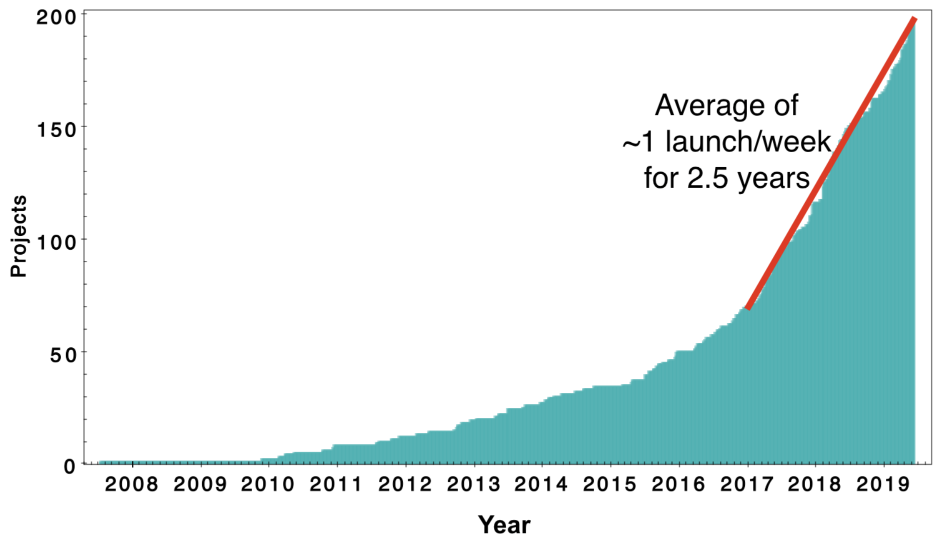

Figure 1: Zooniverse projects launched per year, 2007-2019.

While some teams choose to launch their project publicly through the Zooniverse (after a review process including Zooniverse volunteers providing feedback and improvement suggestions through a standard form), other projects remain private for internal use among research teams, community-based projects, etc.

The Project Builder has been transformative, democratizing access to online citizen science as a tool for research. Prior to its development, a typical online citizen science project required 6-18 months of professional web development time to build. Now the build itself takes a few hours. Zooniverse went from launching 3-5 projects a year to launching 26 in 2016, 44 in 2017, and over 50 in 2018. Fig. 1 shows project growth by year, from 2007 to 2019. Approximately one new public project is launched each week.

5 Recommendations

We have argued above that the current success of citizen science for astronomical data analysis will likely continue into the next decade. Projects such as LSST have adopted citizen science as a core part of their strategy for both scientific exploitation of data and for education and public outreach. We make four recommendations which are designed to maximize the return provided by such projects.

Recommendation 1) *Support and encourage the use of shared infrastructure for online citizen science.*

Astronomers put significant effort into the success of citizen science projects such as Galaxy Zoo or Planet Hunters. Data must be prepared, a project designed to produce scientifically useful results, and analysis of the results requires the same careful work that would apply to any analysis method. In addition, experts answer questions and work with volunteers who have adopted the project. Volunteers give their time, abilities and energy in remarkable degrees, and produce

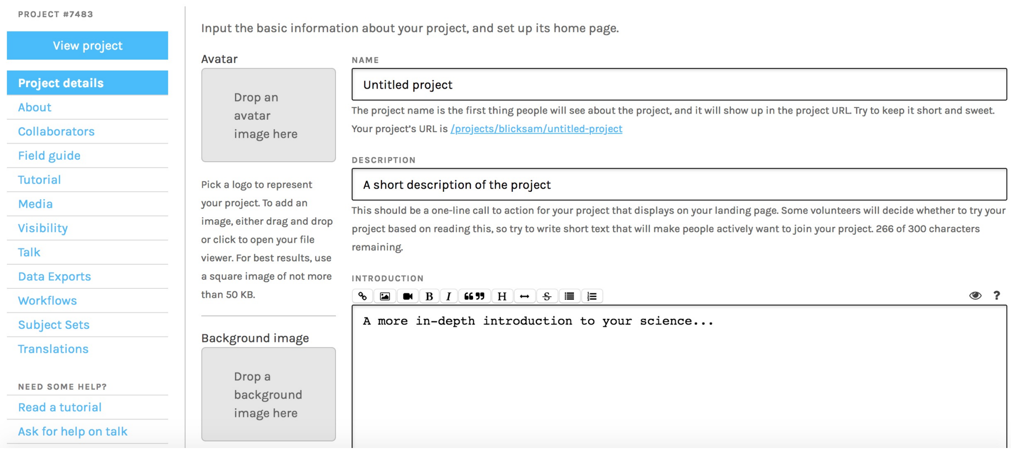

Figure 2: 'Project details' tab of the Zooniverse Project Builder.

remarkable results individually and as a crowd. Yet focusing on this interaction between scientists with data and volunteers who are inspired to participate misses the critical third component provided by the project managers, EPO specialists, web developers, infrastructure engineers, and interaction designers whose work makes projects possible. It also neglects the increasing body of work in the computer science and social science literatures which establishes best practice for designing inclusive and effective citizen science. As a result, creating a project from scratch is likely beyond the capacity of an individual researcher, survey or mission, a situation which is worsened as increasingly sophisticated techniques – for example, dynamic incorporation of machine learning results – leads to greater demands on infrastructure and design.

It is clearly in the interest of the astronomical community for us to share infrastructure. While such sharing starts with the provision of open source code, the Zooniverse model of an installed and open platform, available to researchers for free, has been instrumental in encouraging many to experiment with a citizen science approach to their data analysis. Volunteers benefit from this shared platform approach too, with the barrier to entry on a new project lowered, and the increased possibility of participants forging long-lasting 'careers' as astronomical citizen scientists as they move from one project to another.

The incorporation of the Zooniverse platform into the LSST EPO plans, so that any LSST scientist will be able to easily transfer data to a citizen science project, is an example of how such shared platform design can operate within a single survey. However, there is no need for a barrier between surveys, and we recommend that agencies encourage the missions and surveys they support to contribute to a single shared platform for citizen science wherever possible.

Recommendation 2:) *There should be a default move to make tools and, where plausible, data, developed by missions and large projects available to all.*

The success of citizen scientists in moving beyond classification into serendipitous detection and follow-up has occurred because of the provision of well designed and accessible tools, provided by their creators for all to use. Galaxy Zoo's original choice of SDSS as its dataset enabled volunteers to use the web tools provided by CASJobs, which greatly enhanced the ability of volunteers to investigate the Green Peas. The rich diversity of discoveries from the Planet Hunters project using Kepler data came about because volunteers were able to access the many open tools the exoplanet community has provided. To maximize scientific return, working in the open is key. This is an argument that is often made within professional science; the need and desire of citizen scientists to be full participants makes the requirement especially pressing.

Citizen science projects also produce maximum return where citizen scientists can work alongside their professional peers; we therefore encourage projects to be open from the beginning, rather than delaying releases of data and tools.

Recommendation 3) *Infrastructure requires ongoing support.*

The development of citizen science infrastructure for astronomy has, in the past decade, been provided by a combination of private foundation and agency funding. Both sources emphasize development of novel features or new capacity, rather than maintaining or supporting existing projects. Now online citizen science has matured to the point where a single platform may support projects from many data sources, many of which are intended to be long lived surveys. This situation is increasingly unstable.

The problem is made worse by the adoption of a shared set of tools which will remain stable for many years (Recommendation 1) and by the need to plan for citizen science ahead of the beginning of data collection, as LSST has. Therefore, we recommend support for the infrastructure which supports citizen science at the level of a small archive; e.g., a small version of a NASA data archive such as HEASARC, allowing the team responsible to properly support their community as well as developing new features when needed.

Recommendation 4) *Continue to incorporate explicit language in funding solicitations that encourage use of citizen science in research and education proposals.*

Over the past decade, citizen science has seen increased funding support from National agencies such as the NSF, NASA, NOAO and others. For example, Figure 3 shows the growth in awards made by NSF mentioning citizen science from one in 1998 to over 130 in 2016. However, we note a concerning trend in Figure 3 downwards in the number of grants awarded since 2016. The publication of the Crowdsourcing and Citizen Science Act of 2016 is (perhaps not coincidentally) correlated with the peak of funding seen in Figure 3. Several NSF communiques (e.g., Dear Colleague Letters) were also first sent in 2015 and 2016[3] to engage the research community in support of citizen science.

We recommend that explicit language continue to be issued by agencies to seek proposals related to citizen science. For example, in the 2019 NASA ROSES solicitation, explicit wording encouraging submission of citizen science proposals has been included. We further recommend that agencies continue to hold workshops and other mechanisms to obtain *researcher* community input and buy-in on the evolving uses of citizen science as a research method. It is important that the funding agencies do not succumb to seeing citizen science funding as a "fad" but continue to support the process of integrating it into the basic toolkit of research professionals. To this end, we also recommend that to incentivize the formation of National-scale collaborations for citizen science, solicitations such as those for Science & Technology Centers or Research Collaboration Networks explicitly incorporate language supporting citizen science.

[3]e.g. https://www.nsf.gov/discoveries/disc_summ.jsp?cntn_id=136445 and https://www.nsf.gov/pubs/2016/nsf16059/nsf16059.jsp

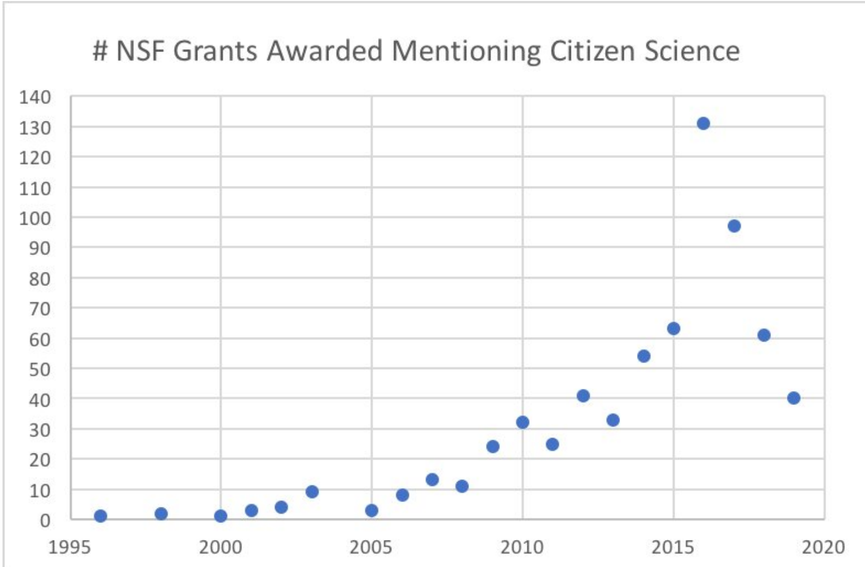

Figure 3: *The number of NSF grants awarded over time based on a search of both current and expired awards with the simple search term "citizen science".*

References

[1] J. K. Banfield, O. I. Wong, K. W. Willett, R. P. Norris, L. Rudnick, S. S. Shabala, B. D. Simmons, C. Snyder, A. Garon, N. Seymour, E. Middelberg, H. Andernach, C. J. Lintott, K. Jacob, A. D. Kapińska, M. Y. Mao, K. L. Masters, M. J. Jarvis, K. Schawinski, E. Paget, R. Simpson, H.-R. Klöckner, S. Bamford, T. Burchell, K. E. Chow, G. Cotter, L. Fortson, I. Heywood, T. W. Jones, S. Kaviraj, Á. R. López-Sánchez, W. P. Maksym, K. Polsterer, K. Borden, R. P. Hollow, and L. Whyte. Radio galaxy zoo: host galaxies and radio morphologies derived from visual inspection. *Monthly Notices of the Royal Astronomical Society*, 453(3):2327–2341, September 2015.

[2] R. Bird, M. K. Daniel, H. Dickinson, Q. Feng, L. Fortson, A. Furniss, J. Jarvis, R. Mukherjee, R. Ong, and I. Sadeh. Muon Hunter: a Zooniverse project. *arXiv e-prints*, page arXiv:1802.08907, Feb 2018.

[3] John H. Debes, Melina Thévenot, Marc J. Kuchner, Adam J. Burgasser, Adam C. Schneider, Aaron M. Meisner, Jonathan Gagné, Jacqueline K. Faherty, Jon M. Rees, Michaela Allen, Dan Caselden, Michael Cushing, John Wisniewski, Katelyn Allers, and and. A 3 gyr white dwarf with warm dust discovered via the backyard worlds: Planet 9 citizen science project. *The Astrophysical Journal*, 872(2):L25, feb 2019.

[4] Hugh Dickinson, Lucy Fortson, Chris Lintott, Claudia Scarlata, Kyle Willett, Steven Bamford, Melanie Beck, Carolin Cardamone, Melanie Galloway, and Brooke Simmons. Galaxy Zoo: Morphological Classification of Galaxy Images from the Illustris Simulation. , 853(2):194, Feb 2018.

[5] Sander Dieleman, Kyle W. Willett, and Joni Dambre. Rotation-invariant convolutional neural networks for galaxy morphology prediction. *Monthly Notices of the Royal Astronomical Society*, 450(2):1441–1459, April 2015.

[6] Carolin Cardamone et al. Galaxy zoo green peas: Discovery of a class of compact extremely star-forming galaxies. *Monthly Notices of the Royal Astronomical Society*, 399(3):1191–1205, 2014.

[7] Chris J. Lintott et al. Galaxy zoo: 'hanny's voorwerp', a quasar light echo? *Monthly Notices of the Royal Astronomical Society*, 399(1):129–140, 2009.

[8] Darryl E. Wright et al. A transient search using combined human and machine classifications. *Monthly Notices of the Royal Astronomical Society*, 472(2):1315–1323, 2017.

[9] Laura Trouille et al. Citizen science frontiers: Efficiency, engagement, and serendipitous discovery with human-machine systems. 2018.

[10] Melanie R. Beck et al. Integrating human and machine intelligence in galaxy morphology classification tasks. *Monthly Notices of the Royal Astronomical Society*, 476(4):5516–5534, 2018.

[11] Michael Zevin et al. Gravity spy: integrating advanced ligo detector characterization, machine learning, and citizen science. *Classical and Quantum Gravity*, 34(6):064003, 2017.

[12] Tabetha Boyajian et al. Planet hunters ix. kic 8462852 – where's the flux? *Monthly Notices of the Royal Astronomical Society*, 457(4):3988–4004, 2016.

[13] L. Fortson, K. Masters, R. Nichol, K.D. Borne, E.M. Edmondson, C. Lintott, J. Raddick, K. Schawinski, and J. Wallin. Galaxy Zoo: Morphological Classification and Citizen Science. In *Advances in Machine Learning and Data Mining for Astronomy, CRC Press, Taylor & Francis Group, Eds.: Michael J. Way, Jeffrey D. Scargle, Kamal M. Ali, Ashok N. Srivastava, p. 213-236*, pages 213–236. 2012.

[14] Lucy Fortson, Darryl Wright, Chris Lintott, and Laura Trouille. Optimizing the Human-Machine Partnership with Zooniverse. *arXiv e-prints*, page arXiv:1809.09738, Sep 2018.

[15] J. E. Geach, A. More, A. Verma, P. J. Marshall, N. Jackson, P.-E. Belles, R. Beswick, E. Baeten, M. Chavez, C. Cornen, B. E. Cox, T. Erben, N. J. Erickson, S. Garrington, P. A. Harrison, K. Harrington, D. H. Hughes, R. J. Ivison, C. Jordan, Y.-T. Lin, A. Leauthaud, C. Lintott, S. Lynn, A. Kapadia, J.-P. Kneib, C. Macmillan, M. Makler, G. Miller, A. Montaña, R. Mujica, T. Muxlow, G. Narayanan, D. O'Briain, T. O'Brien, M. Oguri, E. Paget, M. Parrish, N. P. Ross, E. Rozo, C. E. Rusu, E. S. Rykoff, D. Sanchez-Argüelles, R. Simpson, C. Snyder, F. P. Schloerb, M. Tecza, W.-H. Wang, L. Van Waerbeke, J. Wilcox, M. Viero, G. W. Wilson, M. S. Yun, and M. Zeballos. The Red Radio Ring: a gravitationally lensed hyperluminous infrared radio galaxy at z = 2.553 discovered through the citizen science project SPACE WARPS. *Monthly Notices of the Royal Astronomical Society*, 452:502–510, September 2015.

[16] Kevin C. Harrington, A. Vishwas, A. Weiß, B. Magnelli, L. Grassitelli, M. Zajaček, E. F. Jiménez-Andrade, T. K. D. Leung, F. Bertoldi, and E. Romano-Díaz. The “Red Radio Ring”: Ionised and Molecular Gas in a Starburst/Active Galactic Nuclei at z 2.55. *Monthly Notices of the Royal Astronomical Society*, page 1716, Jul 2019.

[17] William C. Keel, Chris J. Lintott, W. Peter Maksym, Vardha N. Bennert, S. Drew Chojnowski, Alexei Moiseev, Aleksandrina Smirnova, Kevin Schawinski, Lia F. Sartori, and C. Megan Urry. Fading AGN Candidates: AGN Histories and Outflow Signatures. *Astrophysical Journal*, 835(2):256, Feb 2017.

[18] Marc J. Kuchner, Jacqueline K. Faherty, Adam C. Schneider, Aaron M. Meisner, Joseph C. Filippazzo, Jonathan Gagné, Laura Trouille, Steven M. Silverberg, Rosa Castro, Bob Fletcher, Khasan Mokaev, and Tamara Stajic. The first brown dwarf discovered by the backyard worlds: Planet 9 citizen science project. *The Astrophysical Journal*, 841(2):L19, may 2017.

[19] K. Land, A. Slosar, C. Lintott, D. Andreescu, S. Bamford, P. Murray, R. Nichol, M. J. Raddick, K. Schawinski, A. Szalay, D. Thomas, and J. Vandenberg. Galaxy Zoo: the large-scale spin statistics of spiral galaxies in the Sloan Digital Sky Survey. *Monthly Notices of the Royal Astronomical Society*, 388:1686–1692, August 2008.

[20] Chris J. Lintott, Kevin Schawinski, Anže Slosar, Kate Land, Steven Bamford, Daniel Thomas, M. Jordan Raddick, Robert C. Nichol, Alex Szalay, Dan Andreescu, Phil Murray, and Jan Vandenberg. Galaxy Zoo: morphologies derived from visual inspection of galaxies from the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, 389(3):1179–1189, sep 2008.

[21] Philip J. Marshall, Chris J. Lintott, and Leigh N. Fletcher. Ideas for Citizen Science in Astronomy. *Annual Review of Astronomy & Astrophysics*, 53:247–278, Aug 2015.

[22] Office of Science and Technology Policy. Implementation of federal prize and citizen science authority: Fiscal years 2017-20188. *https://www.whitehouse.gov/wp-content/uploads/2019/06/Federal-Prize-and-Citizen-Science-Implementation-FY17-18-Report-June-2019.pdf*, 2019.

[23] J. M. Pasachoff. Halley and his maps of the Total Eclipses of 1715 and 1724. *Journal of Astronomical History and Heritage*, 2:39–54, June 1999.

[24] S. Rappaport, A. Vanderburg, T. Jacobs, D. LaCourse, J. Jenkins, A. Kraus, A. Rizzuto, D. W. Latham, A. Bieryla, and M. Lazarevic. Likely transiting exocomets detected by Kepler. *Monthly Notices of the Royal Astronomical Society*, 474(2):1453–1468, Feb 2018.

[25] S. Rappaport, G. Zhou, A. Vanderburg, A. Mann, M. H. Kristiansen, K. Oláh, T. L. Jacobs, E. Newton, M. R. Omohundro, and D. LaCourse. Deep long asymmetric occultation in EPIC 204376071. *Monthly Notices of the Royal Astronomical Society*, 485(2):2681–2693, May 2019.

[26] B. D. Simmons, Chris Lintott, Kyle W. Willett, Karen L. Masters, Jeyhan S. Kartaltepe, Boris Häußler, Sugata Kaviraj, Coleman Krawczyk, S. J. Kruk, Daniel H. McIntosh, R. J. Smethurst, Robert C. Nichol, Claudia Scarlata, Kevin Schawinski, Christopher J. Conselice, Omar Almaini, Henry C. Ferguson, Lucy Fortson, William Hartley, Dale Kocevski, Anton M. Koekemoer, Alice Mortlock, Jeffrey A. Newman, Steven P. Bamford, N. A. Grogin, Ray A. Lucas, Nimish P. Hathi, Elizabeth McGrath, Michael Peth, Janine Pforr, Zachary Rizer, Stijn Wuyts, Guillermo Barro, Eric F. Bell, Marco Castellano, Tomas Dahlen, Avishai Dekel, Jamie Ownsworth, Sandra M. Faber, Steven L. Finkelstein, Adriano Fontana, Audrey Galametz, Ruth Grützbauch, David Koo, Jennifer Lotz, Bahram Mobasher, Mark Mozena, Mara Salvato, and Tommy Wiklind. Galaxy zoo: quantitative visual morphological classifications for 48 000 galaxies from CANDELS. *Monthly Notices of the Royal Astronomical Society*, 464(4):4420–4447, October 2016.

[27] A. M. Smith, S. Lynn, M. Sullivan, C. J. Lintott, P. E. Nugent, J. Botyanszki, M. Kasliwal, R. Quimby, S. P. Bamford, L. F. Fortson, K. Schawinski, I. Hook, S. Blake, P. Podsiadlowski, J. Jönsson, A. Gal-Yam, I. Arcavi, D. A. Howell, J. S. Bloom, J. Jacobsen, S. R. Kulkarni, N. M. Law, E. O. Ofek, and R. Walters. Galaxy Zoo Supernovae. *Monthly Notices of the Royal Astronomical Society*, 412(2):1309–1319, feb 2011.

[28] M.c.P. Straub. Giving citizen scientists a chance: A study of volunteer-led scientific discovery. *Citizen Science: Theory and Practice*, 1, may 2016.

[29] Mike Walmsley, Lewis Smith, Chris Lintott, Yarin Gal, Steven Bamford, Hugh Dickinson, Lucy Fortson, Sandor Kruk, Karen Masters, and Claudia Scarlata. Galaxy Zoo: Probabilistic Morphology through Bayesian CNNs and Active Learning. *arXiv e-prints*, page arXiv:1905.07424, May 2019.

[30] David Watson and Luciano Floridi. Crowdsourced science: sociotechnical epistemology in the e-research paradigm. *Synthese*, 195(2):741–764, Feb 2018.

[31] Kyle W. Willett, Melanie A. Galloway, Steven P. Bamford, Chris J. Lintott, Karen L. Masters, Claudia Scarlata, B. D. Simmons, Melanie Beck, Carolin N. Cardamone, Edmond Cheung, Edward M. Edmondson, Lucy F. Fortson, Roger L. Griffith, Boris Häußler, Anna Han, Ross Hart, Thomas Melvin, Michael Parrish, Kevin Schawinski, R. J. Smethurst, and Arfon M. Smith. Galaxy zoo: morphological classifications for 120 000 galaxies inHSTlegacy imaging. *Monthly Notices of the Royal Astronomical Society*, 464(4):4176–4203, October 2016.

[32] Kyle W. Willett, Chris J. Lintott, Steven P. Bamford, Karen L. Masters, Brooke D. Simmons, Kevin R. V. Casteels, Edward M. Edmondson, Lucy F. Fortson, Sugata Kaviraj, William C. Keel, Thomas Melvin, Robert C. Nichol, M. Jordan Raddick, Kevin Schawinski, Robert J. Simpson, Ramin A. Skibba, Arfon M. Smith, and Daniel Thomas. Galaxy zoo 2: detailed morphological classifications for 304 122 galaxies from the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, 435(4):2835–2860, September 2013.

[33] M. Willi, R.T. Pitman, A.W. Cardoso, C. Locke, A. Swanson, A. Boyer, M. Veldthuis, and L. Fortson. Identifying animal species in camera trap images using deep learning and citizen science. *submitted to Methods in Ecology and Evolution*, 2018.