

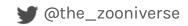
A brief Tour of the Zooniverse: How Crowdsourcing Science is Solving Big Data Problems in Research

Dr. Lucy Fortson
University of Minnesota

Solar Physics Webinar Monthly Speaker Series December 17, 2021





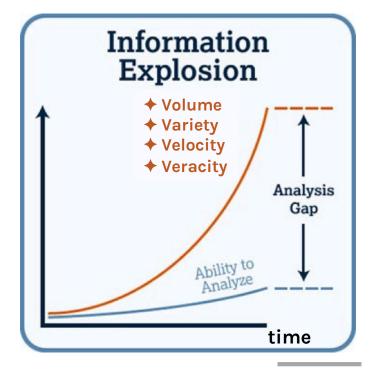






Artificial Intelligence only as good as its training.

The Data Deluge







VS



Which is better at spotting the tiger?







Which is better at spotting the tiger?

Cognitive Surplus?

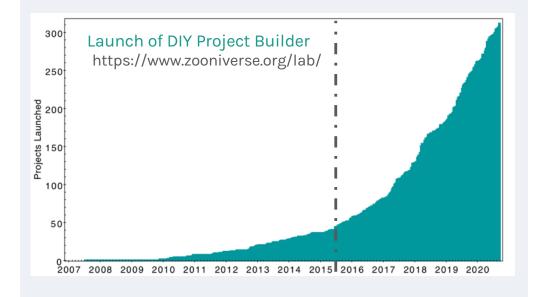
128 years every hour!!

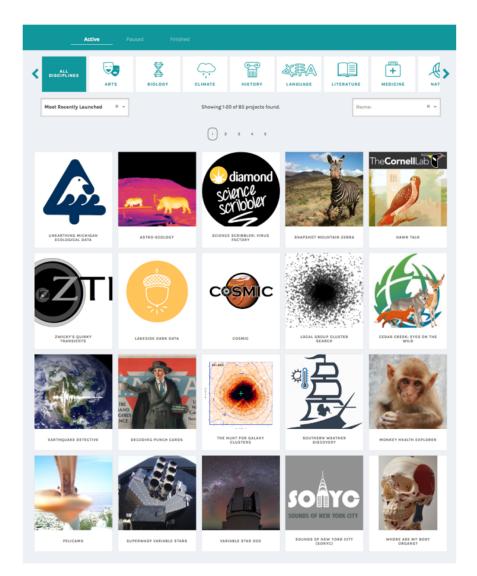




Zooniverse.org

- Started with Galaxy Zoo in 2007
- ~2.5 million volunteers worldwide
- 641,078,369 classifications (as of last night)
- 300+ projects
- 300+ peer reviewed papers





Space Science Projects - active - finished - paused **ZOØNIVERSE** ASAS SN CATALINA OUTER PLANET HUNTERS SPACE FLUFF MUON HUNTER PLANET FOUR: HUBBLE'S HOT MILKY WAY PLANET PATROL CITIZEN ASAS-SN AI4MARS CLASSIC STARS PROJECT FACTORY GALAXY ZOO: HUBBLE ASTEROID EUCLID -MUON HUNTERS SUPERNOVA POPPIN' GALAXY COMET HUNTERS PLANET FOUR: TERRAINS DISK DETECTIVE SPIRAL GRAPH AURORA ZOO PLANET FOUR CLUMP SCOUT CHALLENGE THE HUNTERS ZWICKY'S QUIRKY LOCAL GROUP THE HUNT FOR COSMIC RADIO GALAXY ZOO: LOFAR GALAXY ZOO MOBILE MAPPING HISTORIC SKIES GALAXY BUILDER GALAXY ZOO: 3D ASTEROID ZOO TRANSIENTS CLUSTER SEARCH GALAXY CLUSTERS SUPERWASP VARIABLE STARS PLANET HUNTERS ASTRONOMY STORMWATCH II GALAXY ZOO SUNSPOTTER DISK DETECTIVE RADIO GALAXY PLANET FROM REWIND OUROBOROS

PLANET HUNTERS

SETI LIVE

ANDROMEDA

PROJECT

SOLAR

GRAVITY SPY

WORLDS: PLANET 9

Skv Mapper SUPERNOVA

SIGHTING

NURSERIES

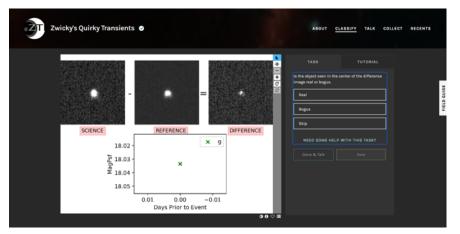
PLANET 9

AGENT NEO

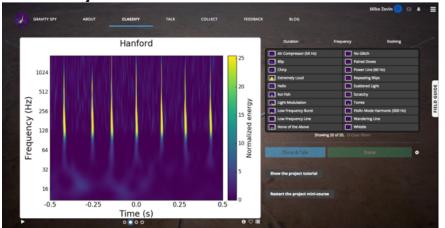
MOON ZOO

Task Types ZOØNIVERSE

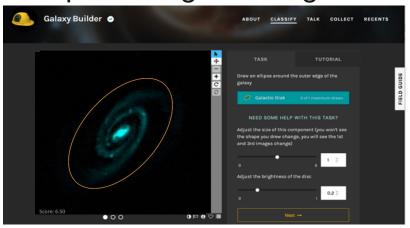
Decision trees



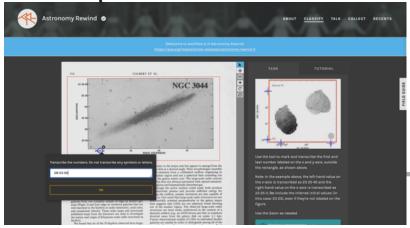
"Survey" or "filter" tool



Multiple marking & drawing tools



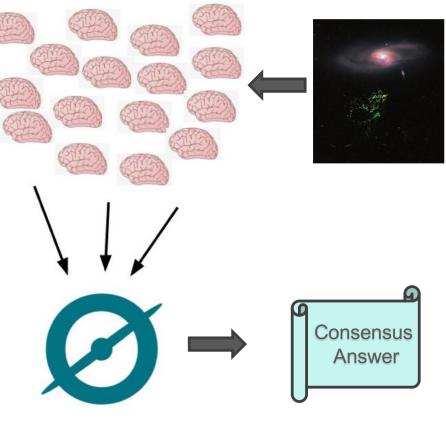
Transcription and annotation tools



How does Zooniverse work?

ZOØNIVERSE

- Volunteers classify (assess data) independently
- Brute force retirement*: between 3 and 80 classifications per image/video file (aka subject)
- Responses are aggregated for consensus
- Raw and consensus data are made available to researchers (and, eventually, open to the public)
- Volunteers interact with researchers on Talk boards, blog posts, social media



*Can use volunteer skill for more sophisticated retirement

- Launched 2007 with 1 million SDSS galaxies
- ~40 million classifications by nearly 150,000 users
- Roughly 3.3 continuous person-years!

Galaxy Zoo solved the intermediate big data issue of not enough "experts" to produce morphological catalogs from surveys on the scale of SDSS.



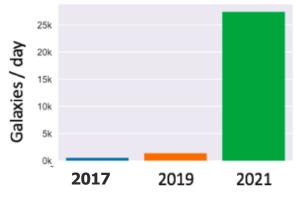
Used brute force retirement of subjects ~40 classifications per galaxy

Increasing Overall System Efficiency

2.0 meter Sloan Digital Sky Survey Telescope



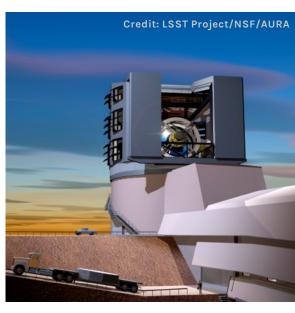
- About 20 Tbytes total data in 10 years
- About 1 million galaxies imaged



But what about the next generation surveys?



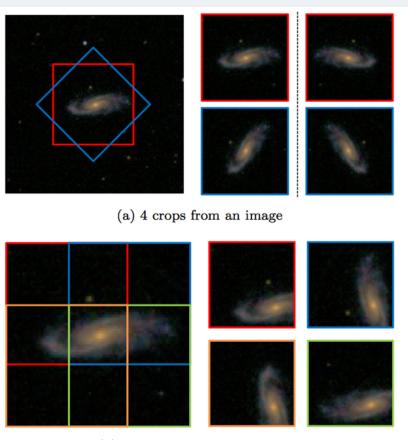
8.4 meter Large Synoptic Survey Telescope

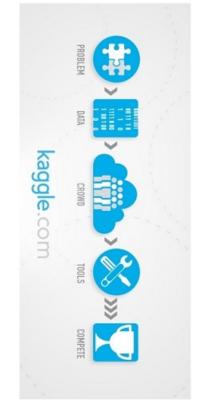


- About 15 Tbytes total data PER NIGHT
- About 50 Petabytes data in 10 years
- About 20 billion galaxies imaged

Training the Machines

Increasing Overall System Efficiency





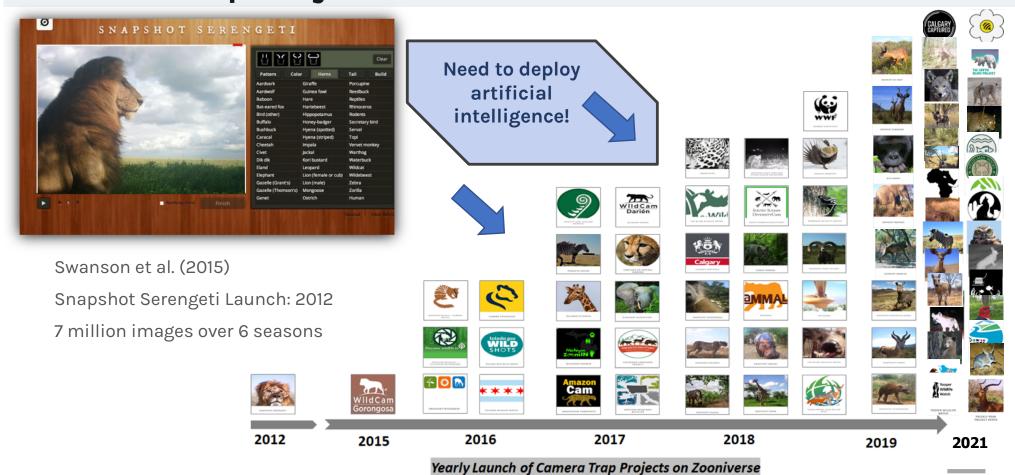
Dieleman et al. arXiv: 1503.07077

(b) 4 viewpoints from each crop

Machines require large sets of images to train them correctly ...

Camera Trap Projects

Deploying Artificial Intelligence



All machines need is just large training sets... ????

Training the Machines

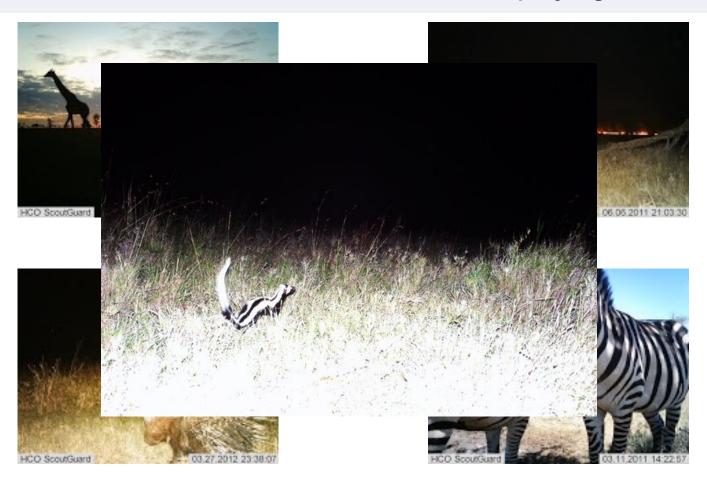
Deploying Artificial Intelligence



Getty Images

The Zorilla Problem

Deploying Artificial Intelligence

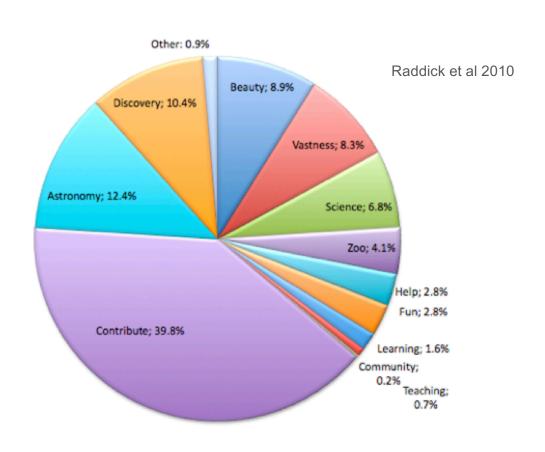


<u>SnapshotSerengeti.org</u>



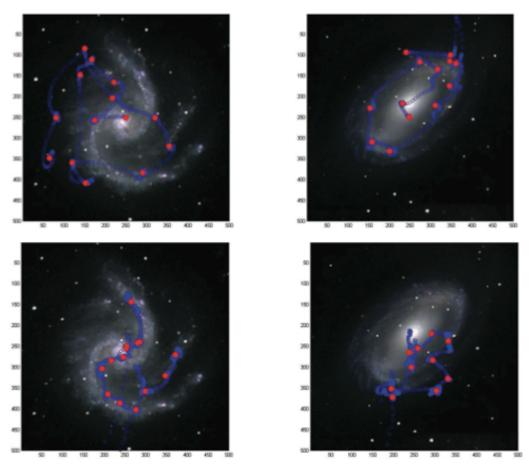
Machine Learning Needs Citizen Science

"Contribute to Research" most common motivation to participate in Galaxy Zoo. Wow!



Machine Learning Needs Citizen Science

Expert vs Non-expert classifier

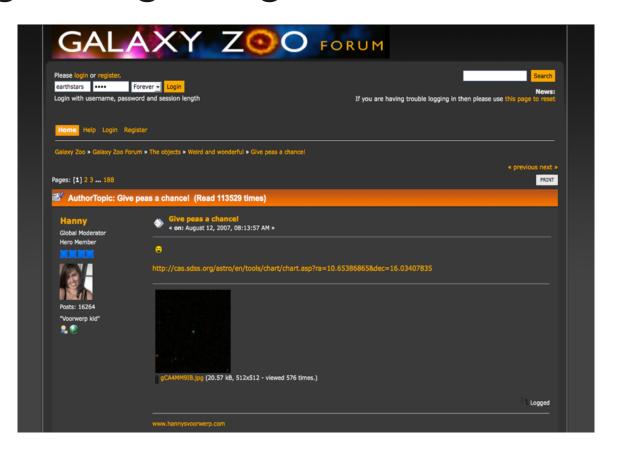


Machine Learning Needs Citizen Science



Understanding Human Classifiers Machine Learning Needs Citizen Science

Posting "strange things" to a discussion board



Machine Learning Needs Citizen Science

DR7

Print

Cross-identifications

catalog delta propermotion angle blue red

Searching for explanations...



0.996203

Humans Make Serendipitous Discoveries

Machine Learning Needs Citizen Science

Galaxy Zoo Green Peas

Entirely new class of galaxy discovered by volunteers!

587732134315425958

J130128.32+510451.2

587732156853846376

J080816.9+281431.1

588013384341913605

587732152555864324

J074758+233632.8

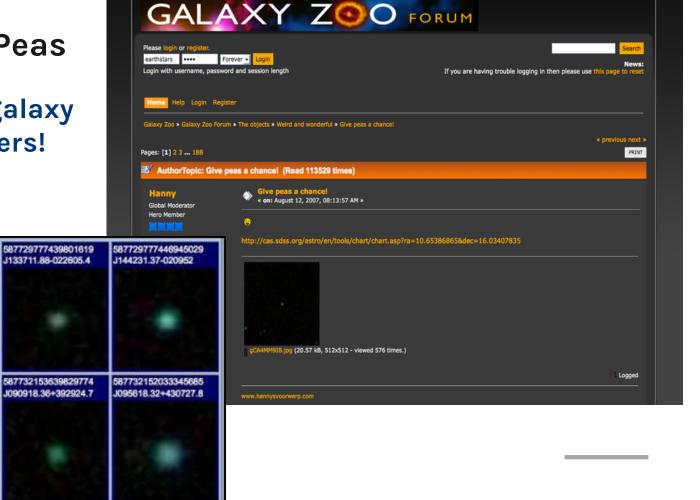
J092600.4+442736.1

588013384336998734

J083833.52+374216.4

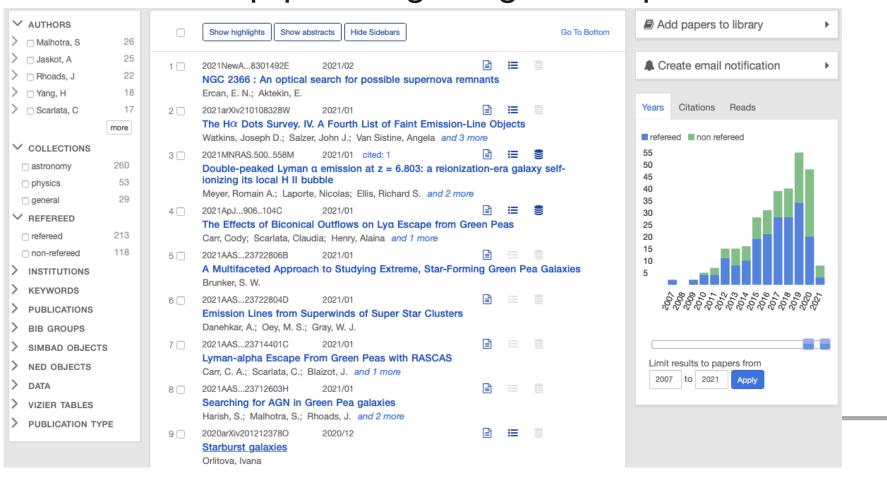
587729970180522426

J150728.4-023351.2



Machine Learning Needs Citizen Science

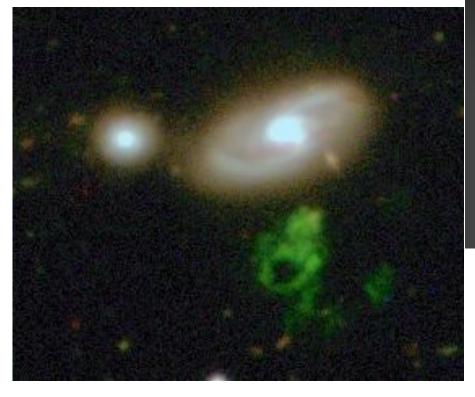
Since 2007, ~330 papers and growing on the topic of Green Peas

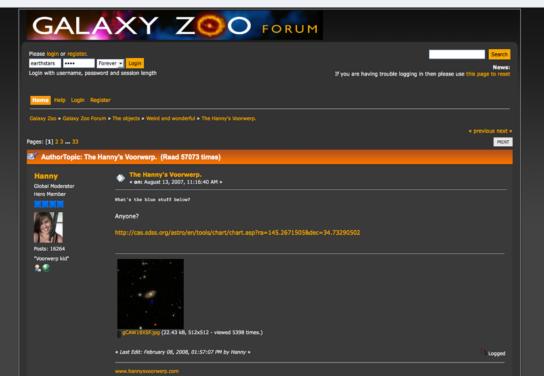


Humans Make Serendipitous Discoveries

Machine Learning Needs Citizen Science

Hanny's Voorwerp





Selected articles from Zooniverse astronomy projects with citizen scientists as co-authors:

Solar Stormwatch:

Davis, ... Baeten (2012)

Barnard, ... Wilkinson, Baeten, Poeffel, Harder (2017)

Space Warps:

Küng, ... Baeten, ... Cornen, Macmillan, ... Wilcox (2015) Geach, ... Baeten, ... Cornen, ... Macmillan, ... Wilcox et al. (2015)

Marshall, ... Wilcox, Baeten, Macmillan, Cornen et al. (2016)

More, ... Baeten, Wilcox, Macmillan, Cornen et al. (2016)

PlanetHunters:

Boyajian, LaCourse, et al. (2016) Schwamb, ... Gagliano, Jek, et al. (2013) Fischer, ... DeFouw, Hajduk, Neal,Nemec, Schuepbach, Zimmermann et al. (2011)

Disk Detective:

Kuchner, ...Biggs, Bosch, Cernohaus, Luca, Hyogo, Wa, Piipuu, Pineiro. (2016)

Radio Galaxy Zoo:

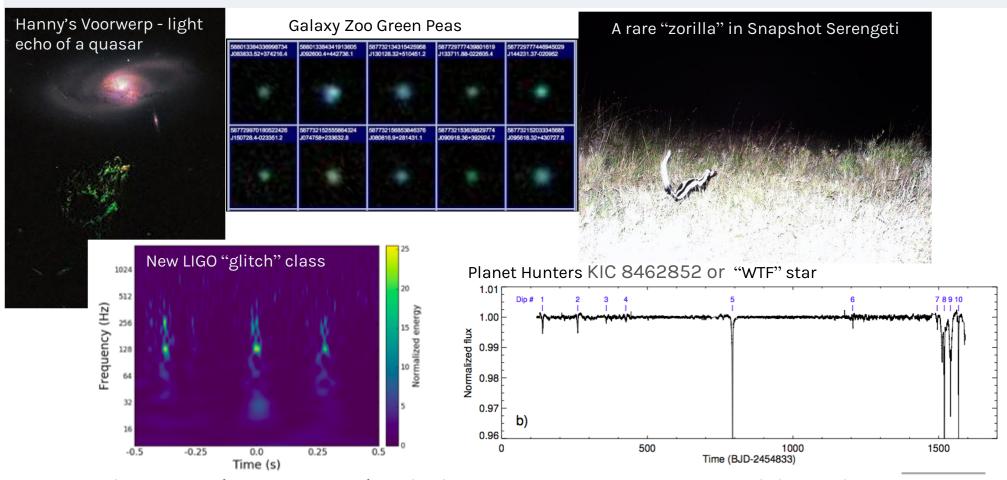
Banfield, ... Matorny, Terentev, et al. (2016)

Galaxy Zoo:

Simmons, ... Jek, et al. (2016) Lintott, ... van Arkel, et al. (2011) Schawinski, ... van Arkel et al. (2014) Keel, ... van Arkel et al. (2014)

Rare and unknown objects

Machine Learning Needs Citizen Science



From primary task (known knowns) to finding "known unknowns" to serendipitous discovery (unknown unknowns) by volunteers.

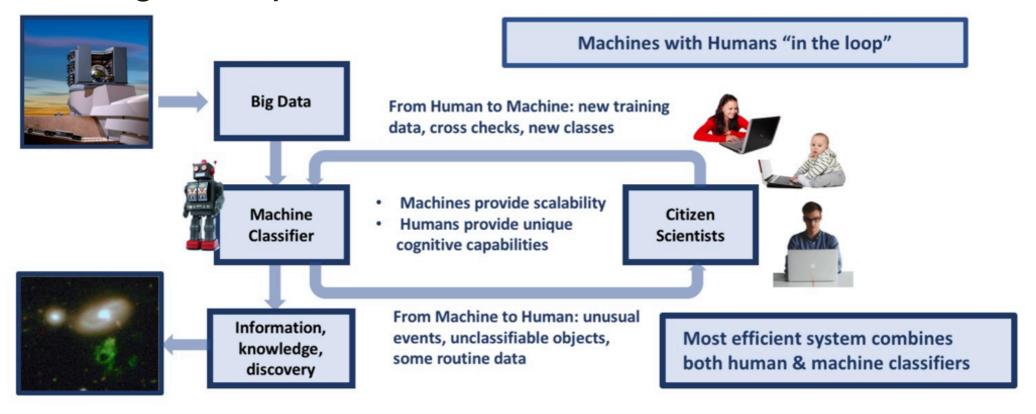
Combining Humans + Machines

ZOØNIVERSE



wallpaperpulse.com – artistic design by Cherif

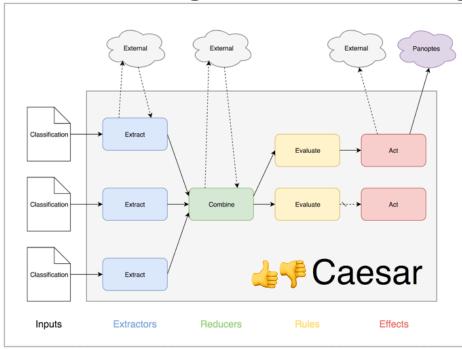
Zooniverse can tackle Big Data by optimally combining machines and humans to quickly get through Big Data while not missing serendipitous discoveries.



Human-in-the-loop Infrastructure

Deploying Artificial Intelligence

Decision engine functionality + Mobile



Made For Mobile

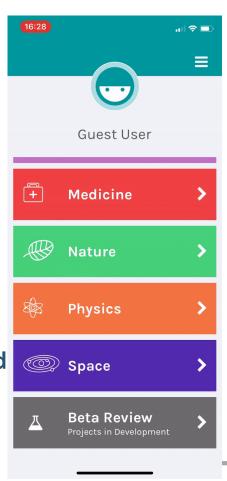
Made For Mobile

General Chilagge

August William Please to Resp Parches
Annual Rig Care

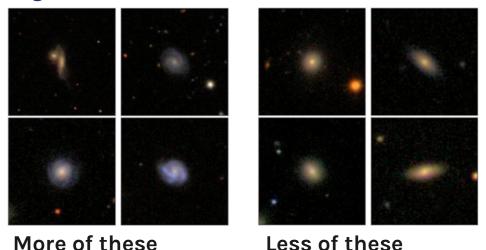
Cascade filtering on mobile: 43% fewer classifications needed

- Early Consensus vs "brute" force retirement
- Dynamic Subject Generation
- User Promotion



Active Learning

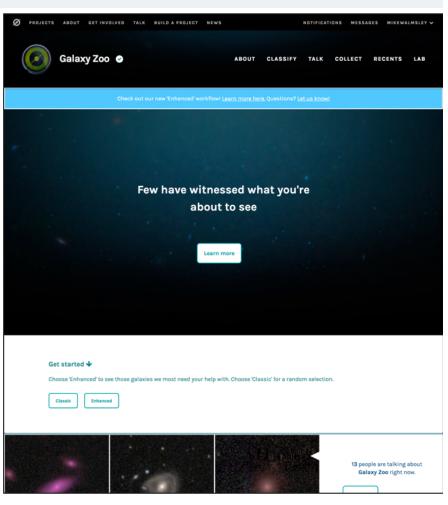
Machine predicts which image, when classified, will give it the most new information



- Model retrains and requests new classifications daily
- New surveys get classified on a timescale of weeks, not years
- Every galaxy seen by at least 3 volunteers

Now live on Galaxy Zoo!

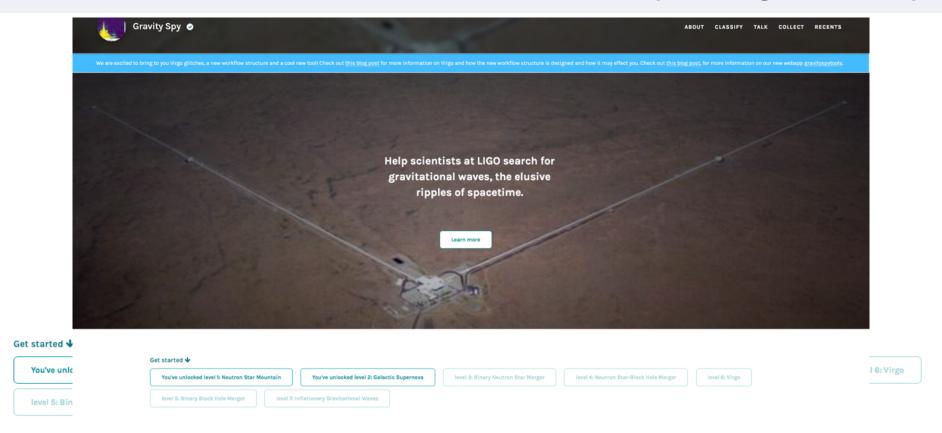
Increasing Overall System Efficiency



Work by Mike Walmsley; Oxford arXiv:1905.07424

Active Learning + Leveling Up Humans

Optimizing for serendipity



Provide volunteers opportunities to "level up" - helps in detecting new types of glitches for LIGO

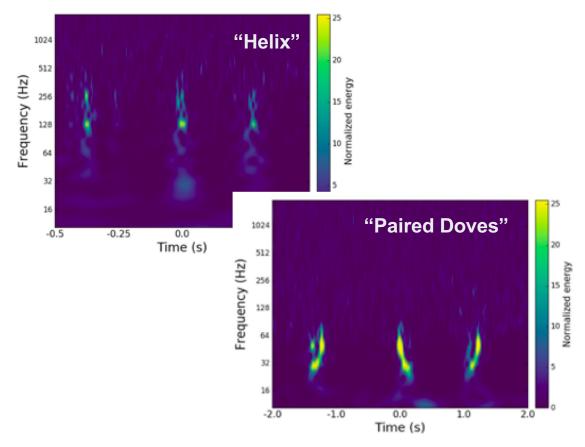
Volunteers use discussion boards to create collections to help identify new glitch classes



January 11th 2017, 9:06 pm

This post is going to be updated many times, first it's a clickable check list (sorry that it's not completely alphabetical). If you know hashtags which are not on the list, please make a comment. If you notice a hashtag what is on the list, and you would like to modify it (because of incorrect spelling, redundancy, permutation of names or for other reason) please feel free to do it in your notes, comments. Thank you for participating in the efforts on making the hashtag system as useful as possible. To be continued!

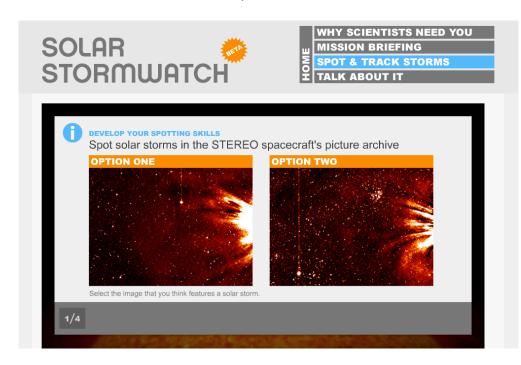
- 1. #aeroline morphology, new. example: Subject 3825220
- 2. #aircompressor official class
- 3. #airplane same as andes
- 4. #also
- 5. #andes same as airplane
- 6. #angel same as mushroom and lfbtree (LF burst variation)
- 7. #anomaly
- 8. #antichirp
- 9. #amplifiedIfb
- 10. #apples morphology example: Subject 2216664
- 11. #arcs morphology, old. scattered light
- 12. #artefact
- 13. #arrow
- 14. #arrowhead



ZOØNIVERSE

Solar Stormwatch was one of the very first projects on the Zooniverse platform!

Launched Dec 21, 2009





ZOØNIVERSE

Solar Stormwatch was one of the very first projects on the Zooniverse platform!

Led to 7 papers...

Solar Stormwatch (7)



Testing the current paradigm for space weather prediction with heliospheric imagers, Barnard+, 2017



Tracking CMEs using data from the Solar Stormwatch project; observing deflections and other properties, Jones+, 2017



Validation of a priori CME arrival predictions made using real-time heliospheric imager observations, Tucker-Hood+, 2015



Observational Tracking of the 2D Structure of Coronal Mass Ejections Between the Sun and 1 AU, Savani+, 2015



Differences between the CME fronts tracked by an expert, an automated algorithm, and the Solar Stormwatch project, Barnard+, 2015



The Solar Stormwatch CME catalogue: Results from the first space weather citizen science project, Barnard+, 2014



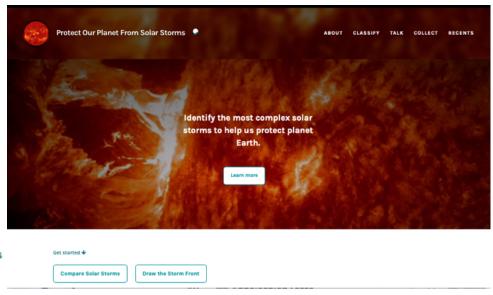
The distribution of interplanetary dust between 0.96 and 1.04 au as inferred from impacts on the STEREO spacecraft observed by the heliospheric imagers, Davis+, 2012

Protect Our Planet From Solar Storms (1)



The Visual Complexity of Coronal Mass Ejections Follows the Solar Cycle, Jones+, 2020

...and a follow-on project from May 2018 to Nov 2020.

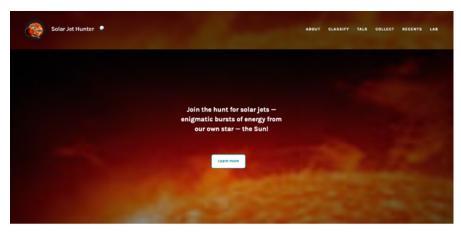


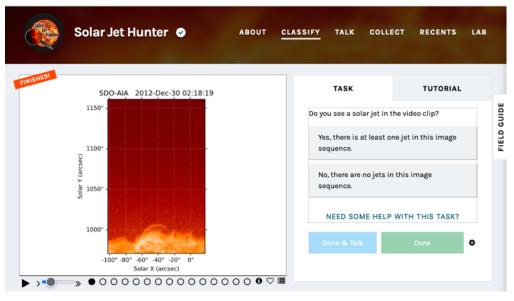
See https://zooniverse.org/publications

ZOØNIVERSE

Newly launched! Solar Jet Hunters with SDO data.

Launched Dec 7, 2021



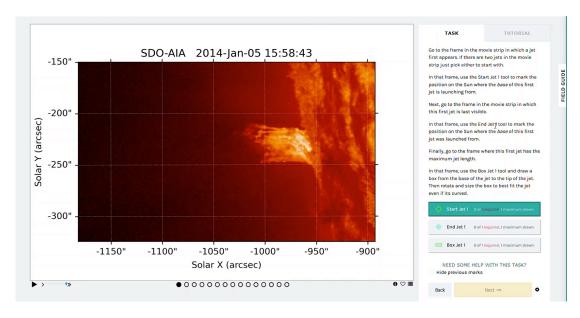


- Took only 2 days for 700 volunteers to classify 4000 clips in data from 2011-2013
- 1200 clips contained at least one solar jet confirmed by at least two people
- First workflow to identify if solar jets are in a 15 frame clip

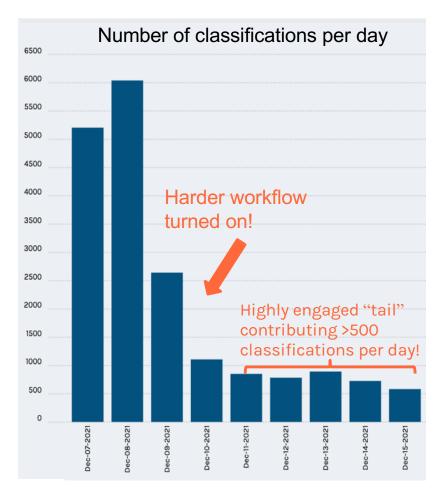
Led by Sophie Musset, Lindsay Glesener, Gregory Fleishman, Navdeep Panesar, et al.

ZOØNIVERSE

Newly launched! Solar Jet Hunters with SDO data.



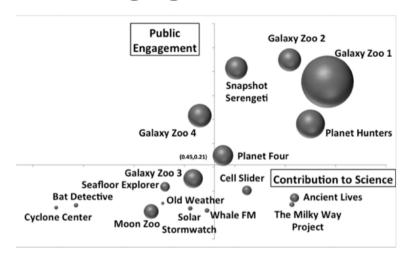
- Second work flow provide location of jet base at start/end of jet and length/width at maximum extent.
- Ultimate goal: develop machine learning algorithm to identify and locate solar jets.



Volunteer Engagement

ZOØNIVERSE

Human Engagement critical to science outcomes



Cox et al 2015

BUT...



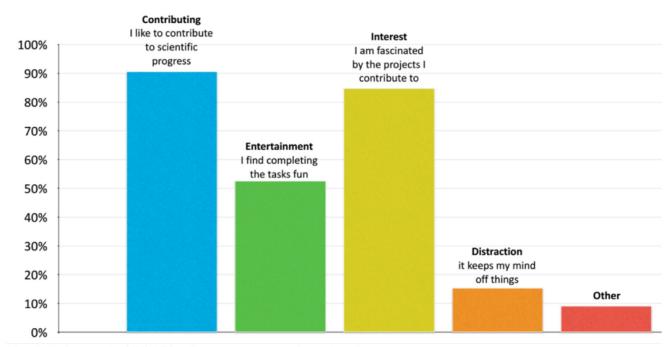
Bowyer et al 2015 HCOMP

When tasks are hard or machines are in the loop, need to be mindful that people are complicated...

When tasks are hard or machines are in the loop, need to be mindful that people are complicated...

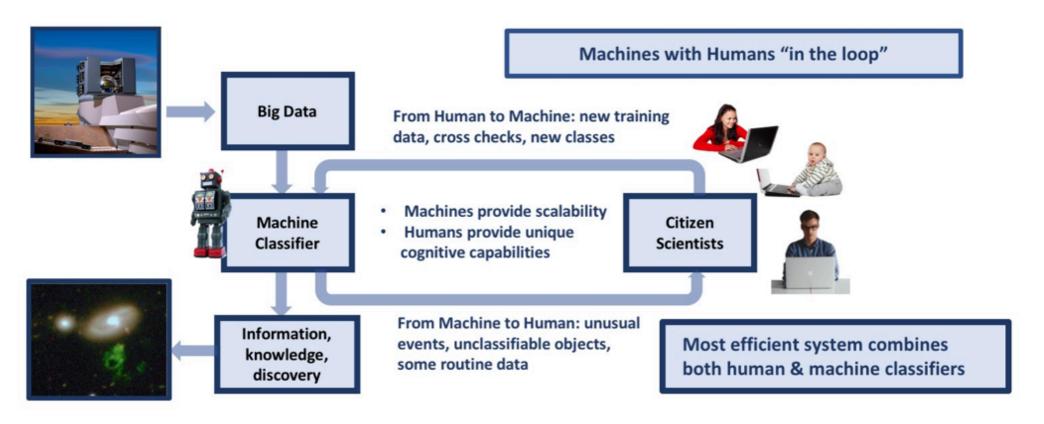
Recall volunteer incentive: contribute to research

- Many will stay through hard projects if you keep them engaged.
- Make sure very clear messaging to volunteers if machines are deployed at all!



Faddick et al. (2013); Crowaton et al. aubmittac

Zooniverse is ready to help close the analysis gap in astrophysics, solar science and more!



Advertisement ZOØNIVERSE

ZOØNIVERSE





Zooniverse Announces NASA Partnership Leveraging World Leading Scientific Expertise

Press Release From: Adler Planetarium Posted: Wednesday, July 22, 2020

Chicago's Adler Planetarium, along with its Zooniverse team and the team from the University of Minnesota, and NASA are pleased to announce a new partnership leveraging world-leading expertise across organizations to advance citizen science capabilities and achievements. Zooniverse is the largest platform for people-powered research in the world. To date, more than two million Zooniverse volunteers have come together to assist professional researchers in over 240 projects, enabling research that would not be possible, or practical, otherwise. Volunteers don't need any specialized background, training, or expertise to participate in any Zooniverse projects, and can contribute to real academic research, on their own computers, at their own convenience.

If you have an idea for a NASA-funded citizen science project apply to the <u>Citizen Science Seed Funding Program:</u>

https://science.nasa.gov/science-news/citizenscience/nasa-funding-available-for-scientists

And thanks to all of our volunteers!!!

Thank you!

lucy@zooniverse.org

y @lucyfortson

LEVERHULME TRUST _____





















