University of Minnesota
Driven to Discover®

ADLER PLANETARIUM

UNIVERSITY OF OXFORD

ZOONIVERSE

Crowdsourcing your training labels with Zooniverse

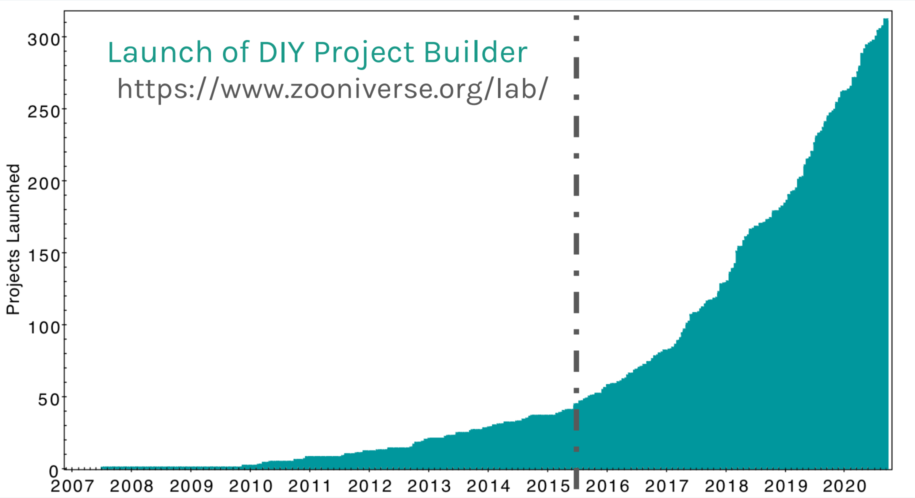
Dr. Lucy Fortson
University of Minnesota

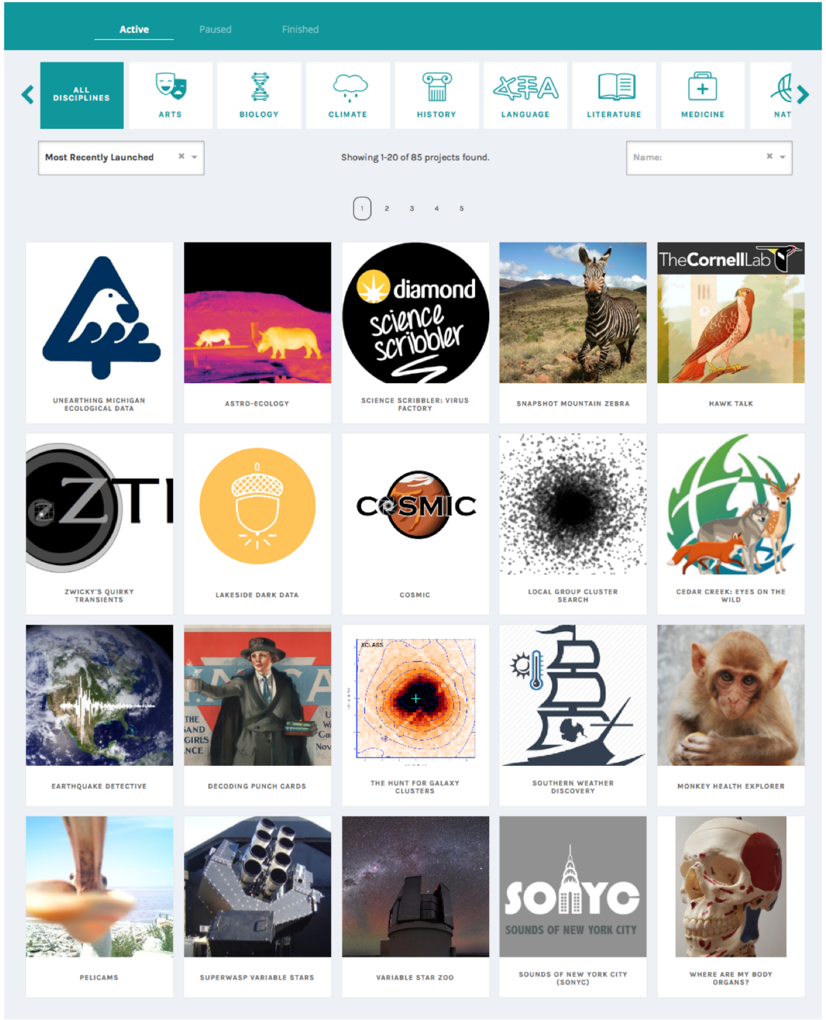
Workshop on Machine Learning for Cosmic-Ray Air Showers
February 3, 2022

zooniverse.org @the_zooniverse @therealzooniverse

Zooniverse.org

- Started with Galaxy Zoo in 2007
- ~2.5 million volunteers worldwide
- 652,069,074 classifications (as of this morning)
- 300+ projects
- 300+ peer reviewed papers (see zooniverse.org/publications)

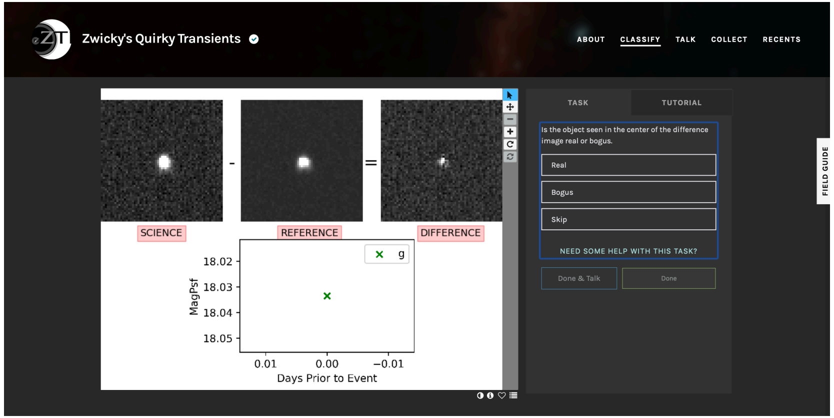




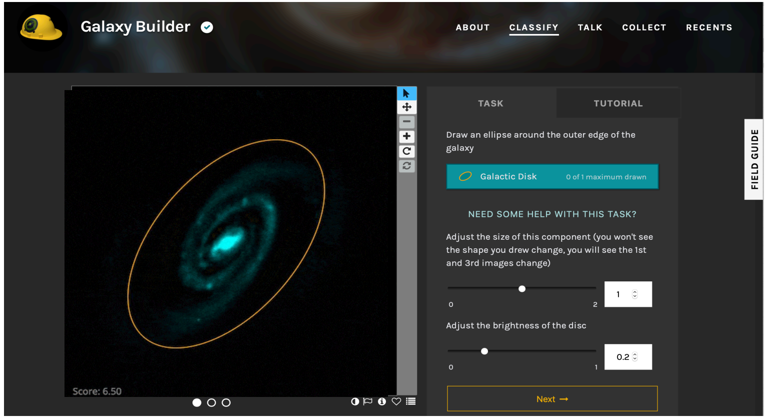
Task Types

ZOONIVERSE

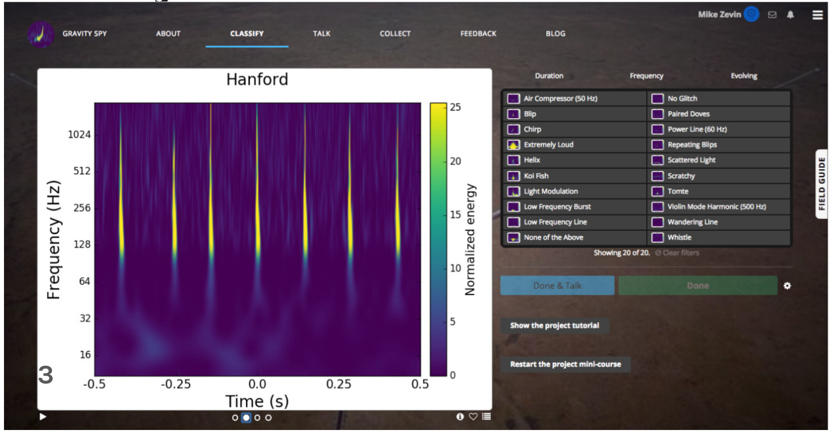
Decision trees



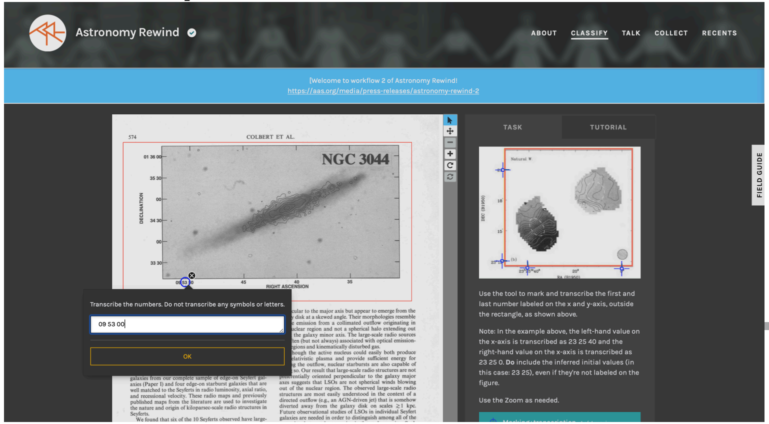
Multiple marking & drawing tools



“Survey” or “filter” tool



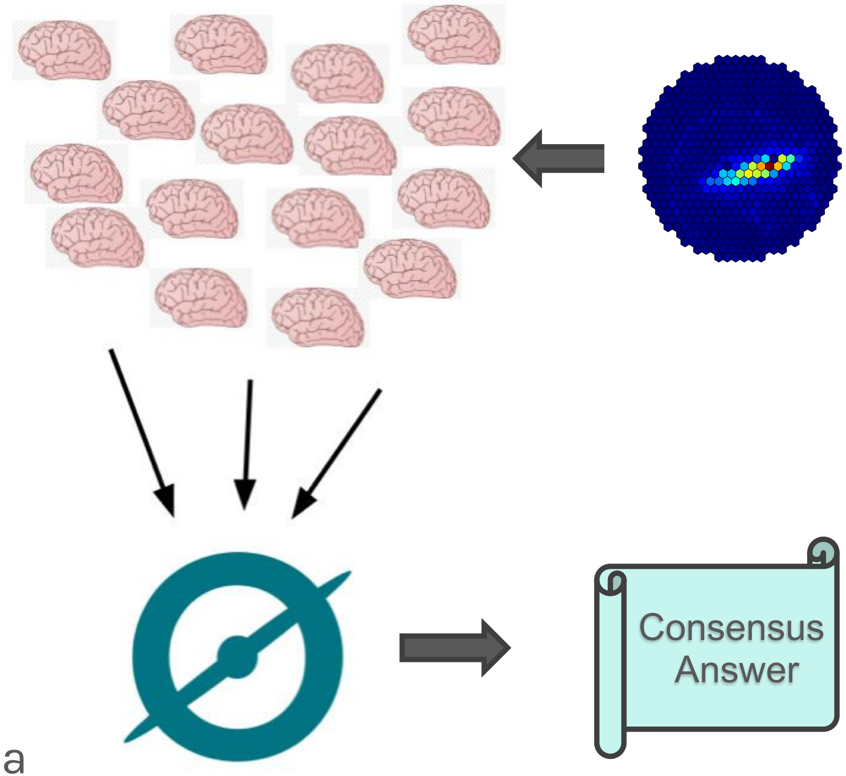
Transcription and annotation tools



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

Citizen science: rapidly growing

Zooniverse Project Builder

Public – intended for promotion by Zooniverse

- 100,000 users signed up for beta testing

Public – intended for promotion by research team with URL

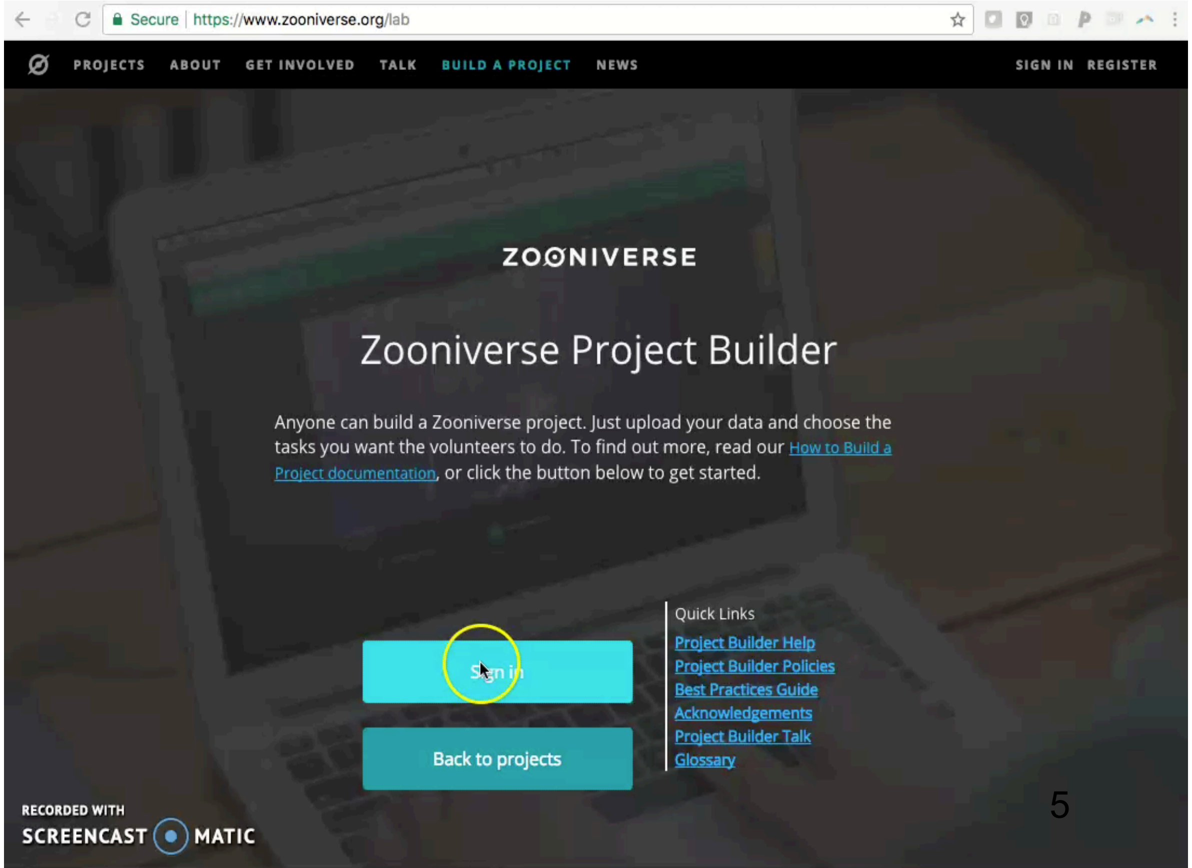
- Ideal for education and small research projects

Private – behind firewall

- Useful for projects requiring human subject review

https://www.zooniverse.org/lab/

Project Builder Demo

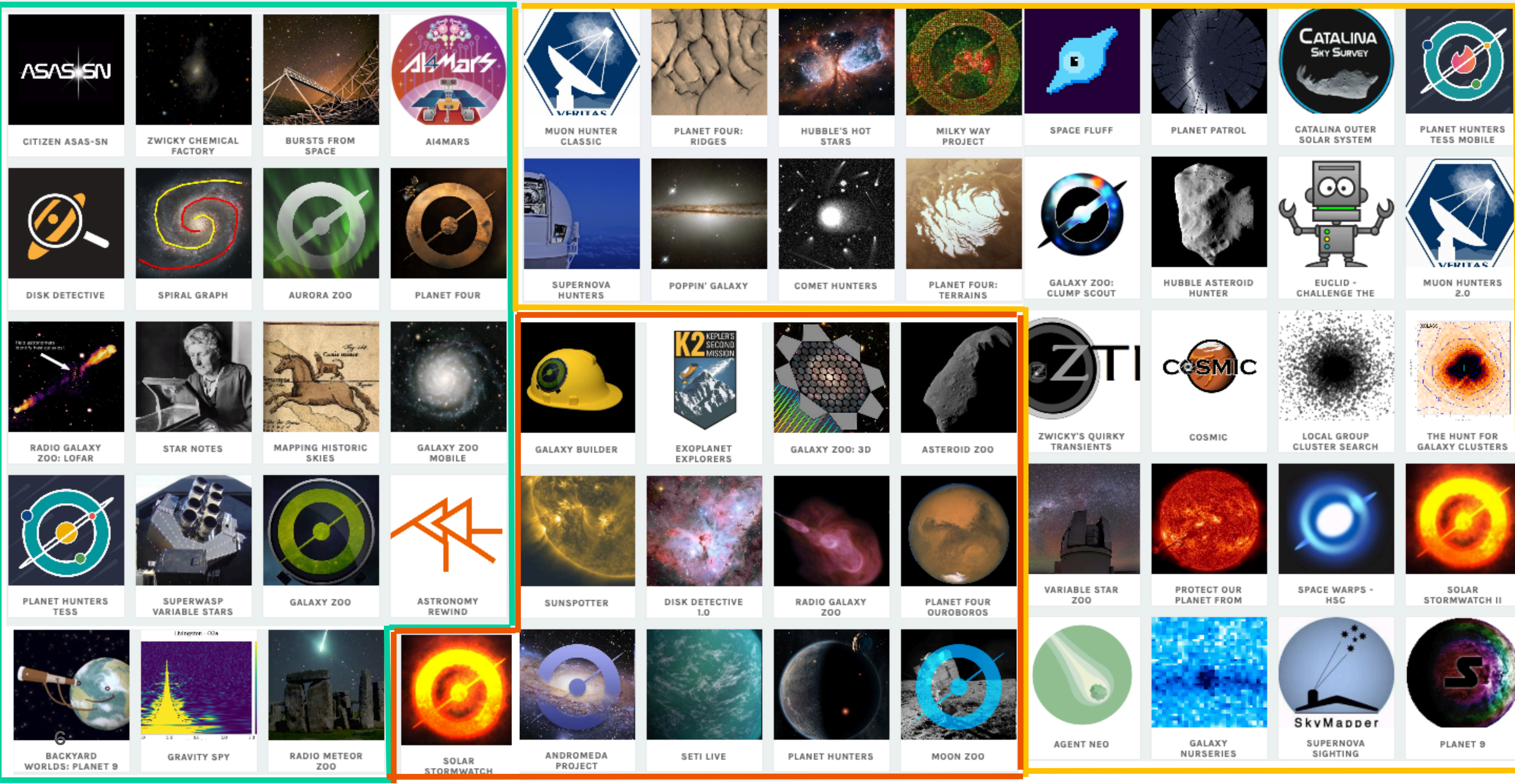


5

Space Science Projects

- active
- finished
- paused

ZOONIVERSE



The Galaxy Zoo Era

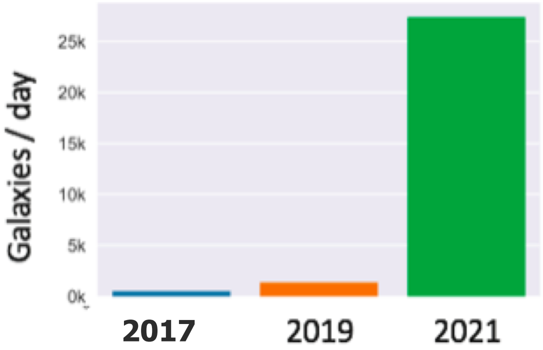
- Launched 2007 with 1 million SDSS galaxies
- ~40 million classifications by nearly 150,000 users
- Roughly 3.3 continuous person-years!

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

2.0 meter Sloan Digital Sky Survey Telescope

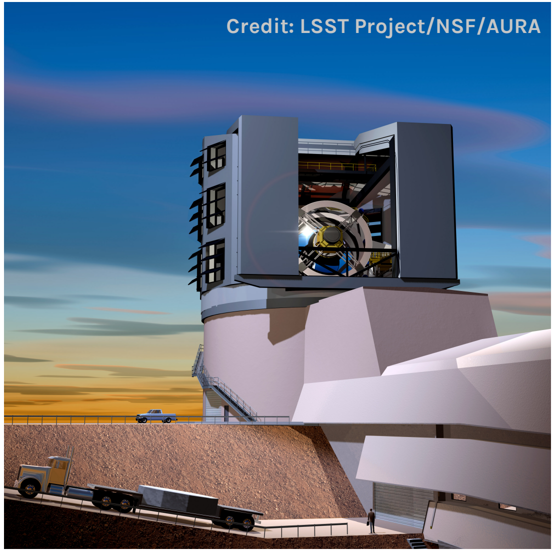
- About 20 Tbytes total data in 10 years
- About 1 million galaxies imaged



But what about the next generation surveys?

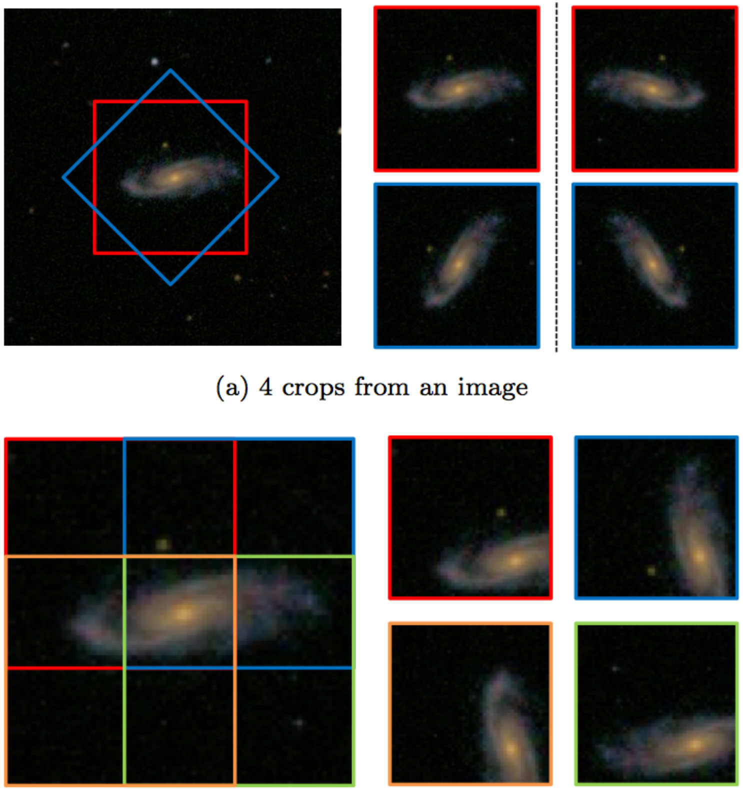
Need to deploy artificial intelligence!

8.4 meter Large Synoptic Survey Telescope



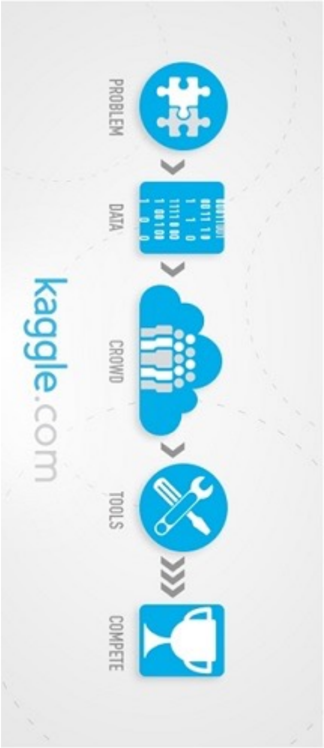
- About 15 Tbytes total data PER NIGHT
- About 50 Petabytes data in 10 years
- About 20 billion galaxies imaged

Increasing Overall System Efficiency



(a) 4 crops from an image

(b) 4 viewpoints from each crop



Dieleman et al.
arXiv: 1503.07077

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

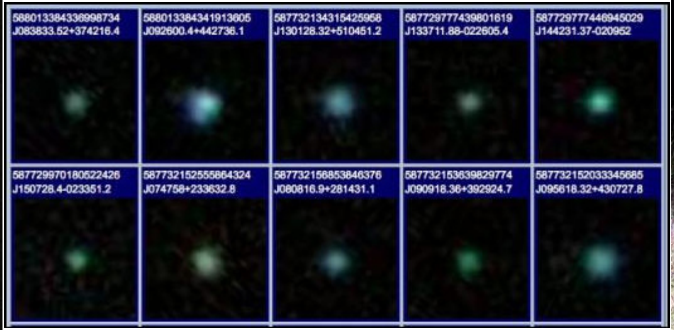
Deploying Artificial Intelligence

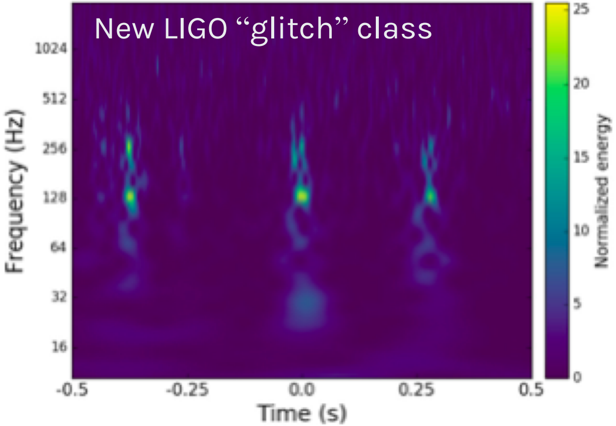
Are we really ready for this?

Getty Images

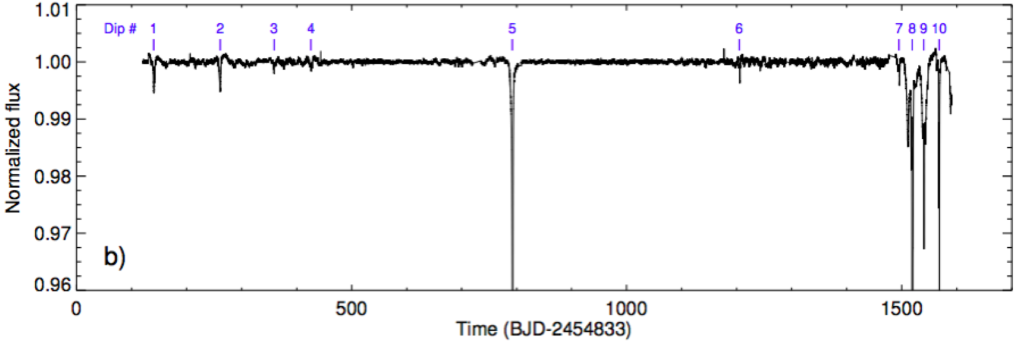
Rare and unknown objects

Galaxy Zoo Green Peas





Planet Hunters KIC 8462852 or "WTF" star



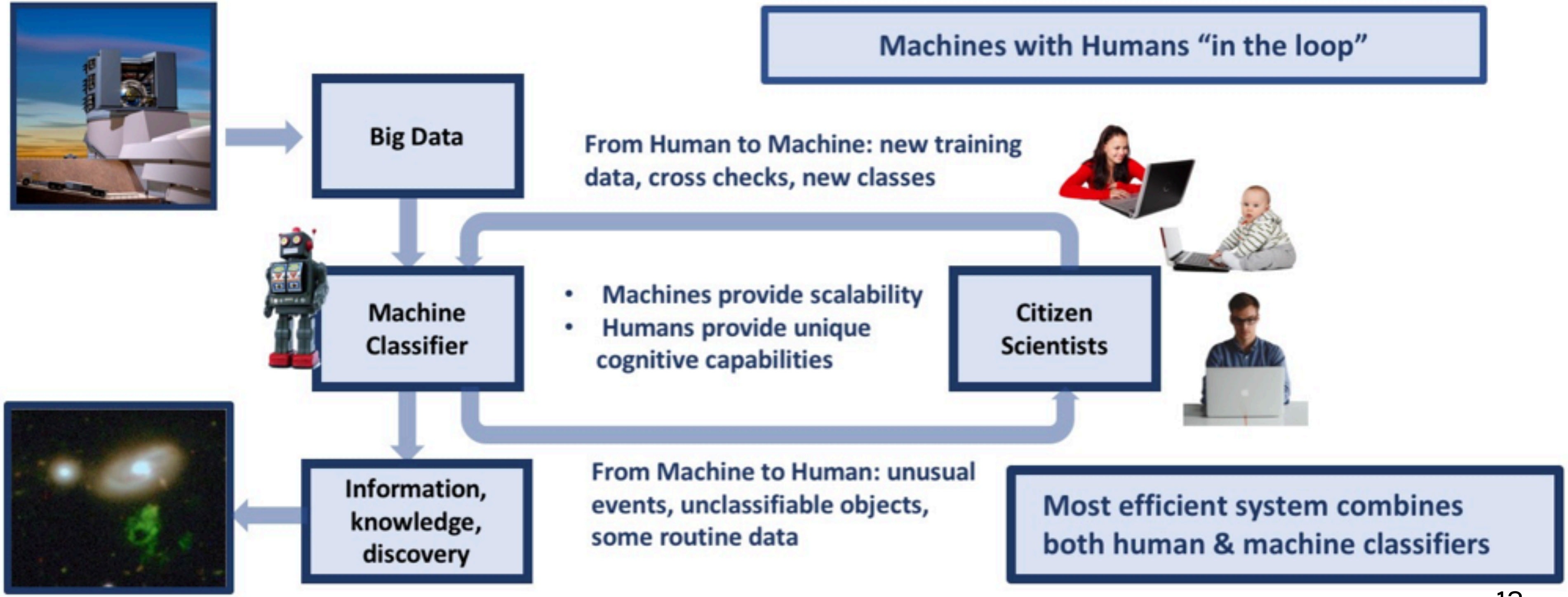
From primary task (known knowns) to finding "known unknowns" to serendipitous discovery (unknown unknowns) by volunteers.

Combining Humans + Machines

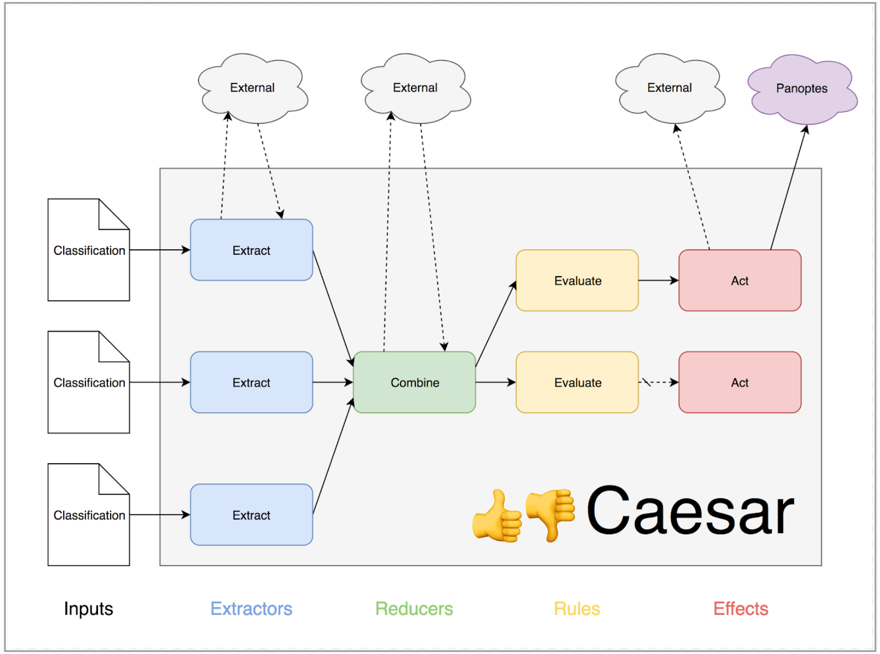
ZOØNIVERSE

wallpaperpulse.com – artistic design by Cherif

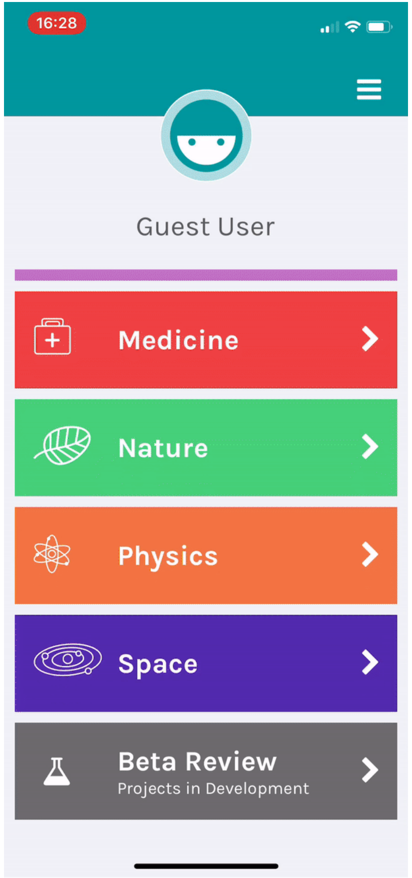
Zooniverse can tackle Big Data by optimally combining machines and humans to quickly get through Big Data while not missing serendipitous discoveries.



Decision engine functionality + Mobile



Cascade filtering on mobile: 43% fewer classifications needed*

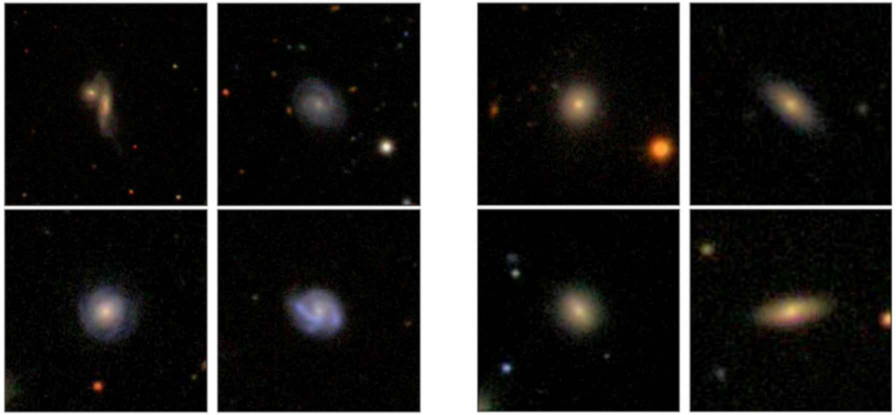


- Early Consensus vs “brute” force retirement
- Dynamic Subject Generation
- Volunteer Promotion based on Skill

*Willi, M. *Methods in Ecology and Evolution*, 2018

Increasing Overall System Efficiency

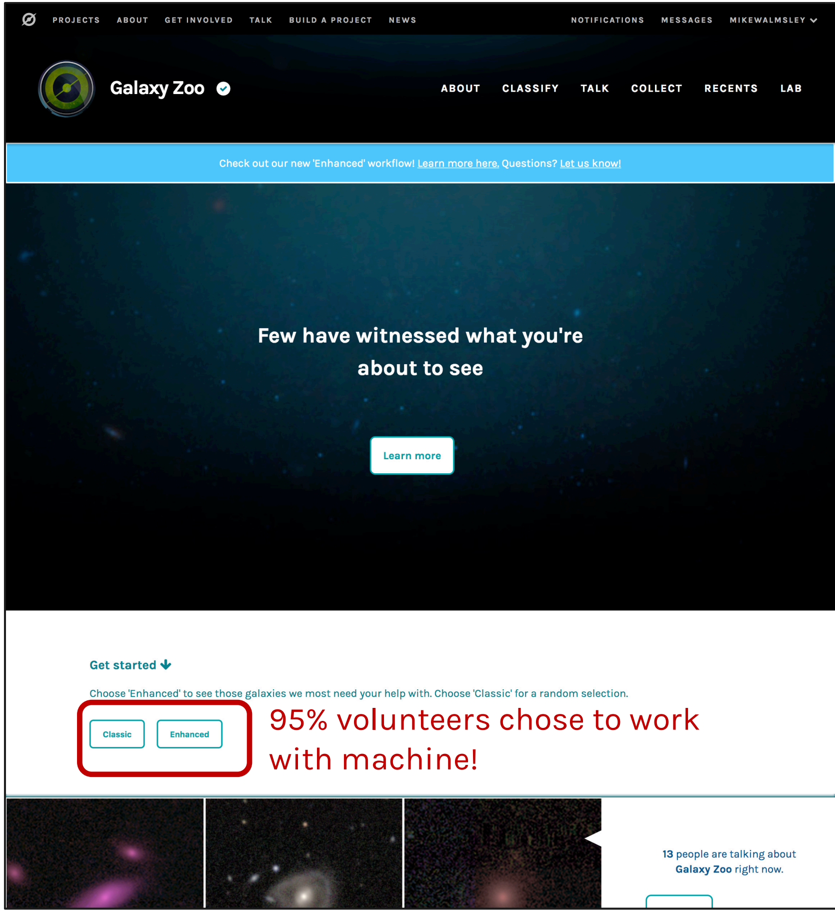
Machine predicts which image, when classified, will give it the most new information



More of these — **Less of these**

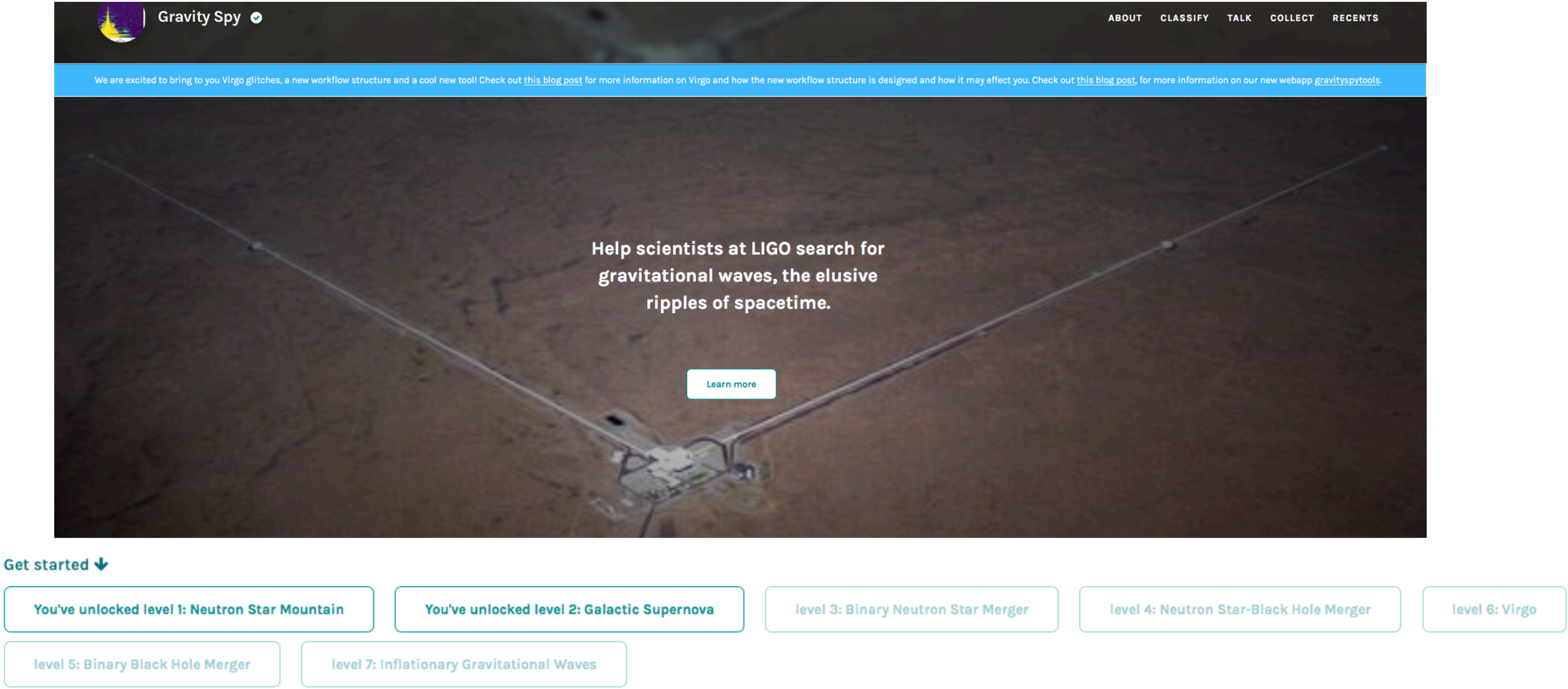
- Model retrains and requests new classifications daily
- New surveys get classified on a timescale of weeks, not years
- Every galaxy seen by at least 3 volunteers

Now live on Galaxy Zoo !



Work by Mike Walmsley; Oxford arXiv:1905.07424

Active Learning + Leveling Up Humans



Provide volunteers opportunities to “level up” – helps in detecting new types of glitches for LIGO

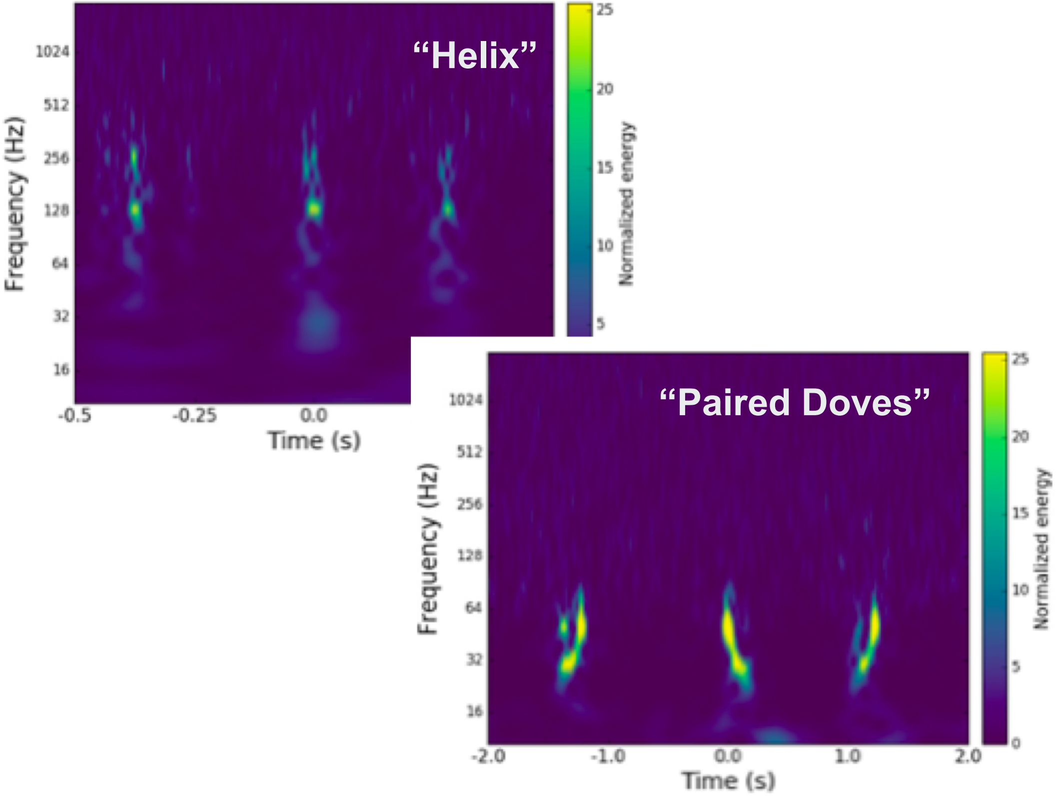
Volunteers use discussion boards to create collections to help identify new glitch classes

EcceruElme
@EcceruElme
MODERATOR

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. #amplifiedlfb
10. #apples morphology example: Subject 2216664
11. #arcs morphology, old. scattered light
12. #artefact
13. #arrow
14. #arrowhead



Some recent machine learning papers using Zooniverse projects ZOØNIVERSE

1. Practical Galaxy Morphology Tools from Deep Supervised Representation Learning. (Walmsley, M. et al, MNRAS, 2022; arXiv:2110.12735)
2. Citizen ASAS-SN Data Release I: Variable Star Classification Using Citizen Science. (Christy, C. T. et al, MNRAS 2022; arXiv:2111.02415v1)
3. Can't we all just get along? Citizen scientists interacting with algorithms. (Ponti, M. et al, Human Computation, 2021; https://doi.org/10.15346/hc.v8i2.128)
4. Transient oscillations in steelpan drums tracked via machine learning. (Hawly, S. et al, J. Acoust. Soc., 2021; https://doi.org/10.1121/10.0008030)
5. Automated detection of surface changes on comet 67P. (Vincent, J-P. et al, EPSC 2021; 10.5194/epsc2021-525)
6. Discovering features in gravitational-wave data through detector characterization, citizen science and machine learning. (Soni, S. et al, Classical and Quantum Gravity, 2021; https://doi.org/10.1088/1361-6382/ac1ccb)
7. Deep learning for automatic segmentation of the nuclear envelope in electron microscopy data, trained with volunteer segmentations. (Spiers, H. et al, Traffic 2021; **https://doi.org/10.1111/tra.12789**)
8. Time-lapse imagery and volunteer classifications from the Zooniverse Penguin Watch project. (Jones, F. M. et al, Nature 2018, https://doi.org/10.1038/sdata.2018.124)

But we can get labels from simulations…

ZOONIVERSE

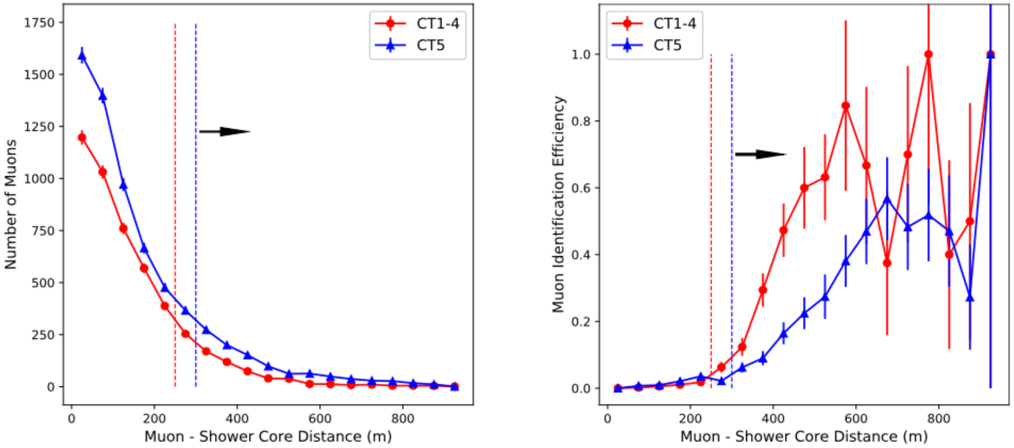
Simulations are only as good as our understanding of physics and our ability to model same.

1. For UHECR energies, potential divergence between real airshower data and sims tuned by LHC data but *extrapolated* from LHC energies
2. For IACT energies, *expect* good agreement as no extrapolation needed. But…hints of divergence in muon distributions between LHC-tuned models and data.

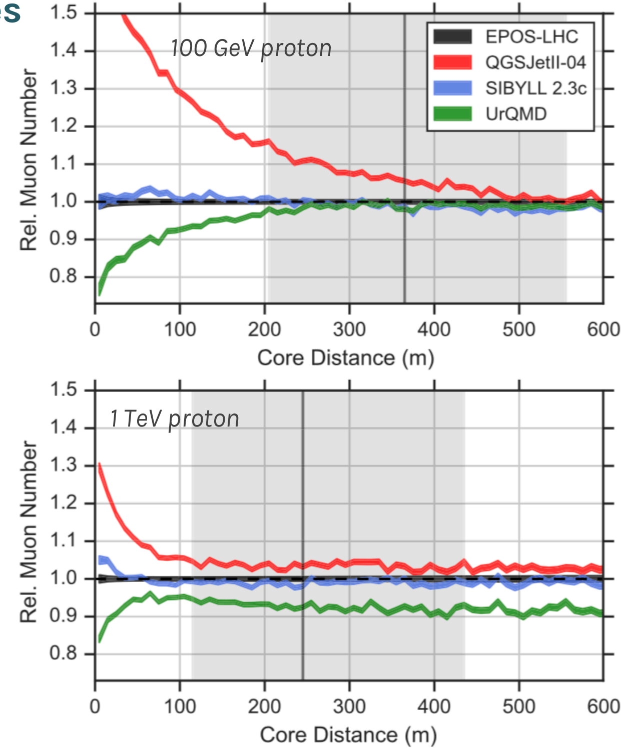
Human labelers can help circumvent these issues!

IACT measurement of "clean" muons difficult at relevant core distances and energies that test models.

H.E.S.S. telescopes reconstruction of muons simulated with QGSJetII-04



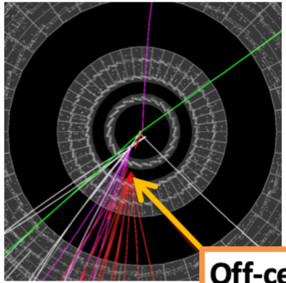
A.M.W. Mitchell, H.P. Dembinski and R.D. Parsons / Astroparticle Physics 111 (2019) 23–34



R.D. Parsons & H. Schoorlemmer/Phys. Rev. D 100, 023010 (2019)

Example projects related to particle and multi-messenger astrophysics

HiggsHunters.org



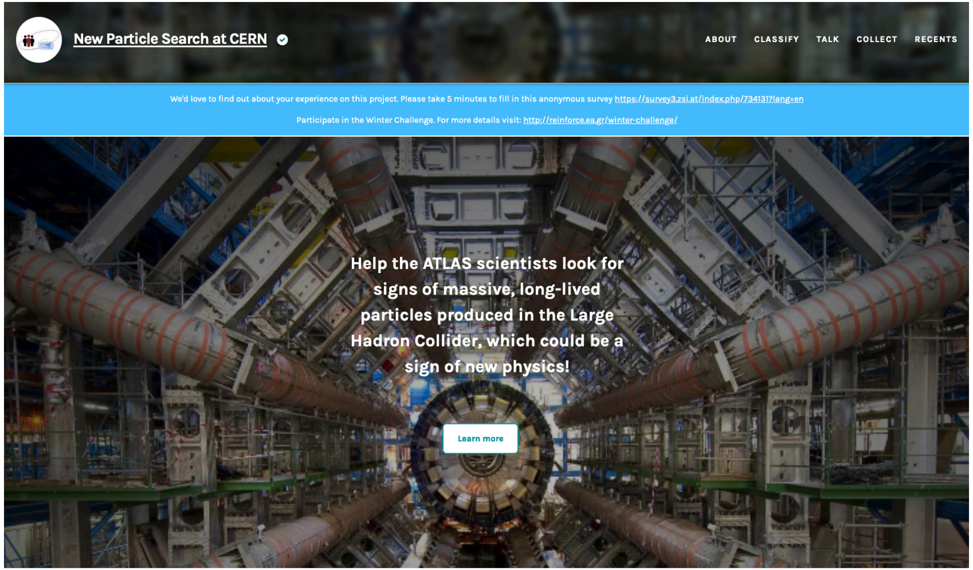
First Zooniverse "particle physics" project:

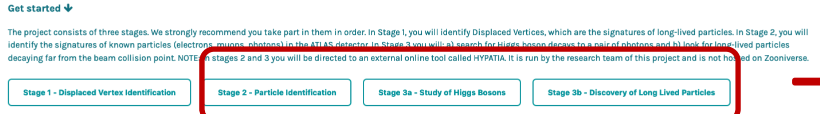
- 2014-2016
- 38,000 citizen scientists
- 39,000 images
- 1,500,000 features
- Barr et al 2016: https://arxiv.org/pdf/1610.02214.pdf

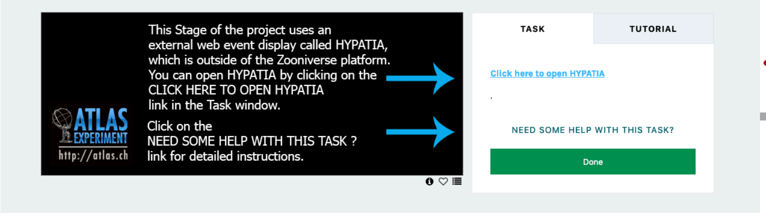
Off-centre vertex

A non-SM boson travels some distance before spraying out tracks

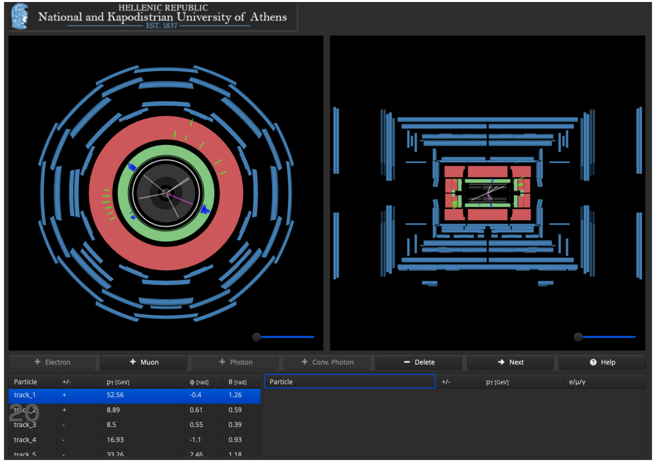
Just launched new particle physics project!





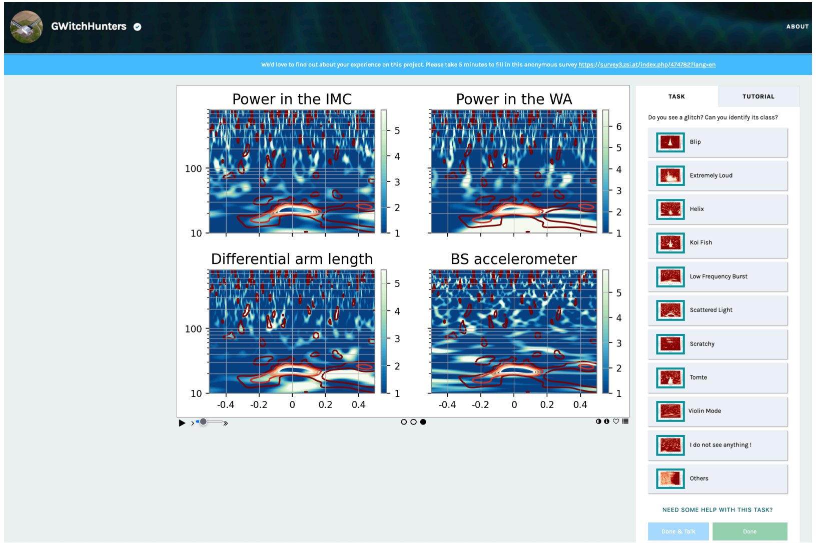


now incorporating ATLAS' own event display

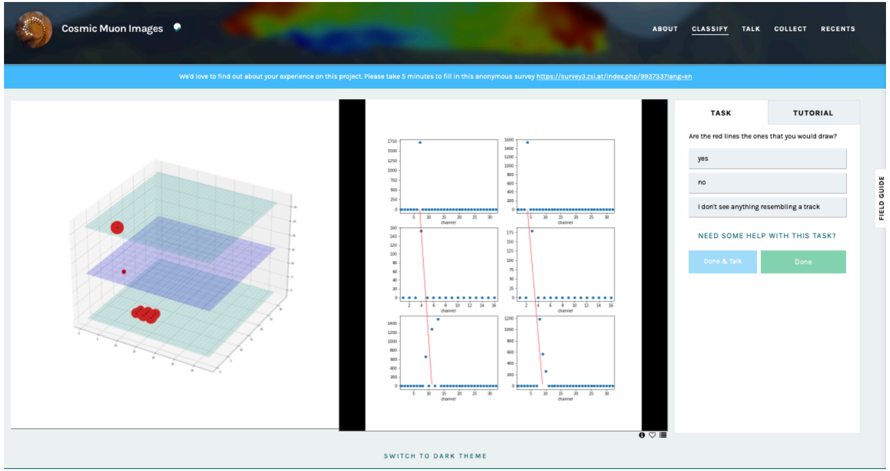


Example projects related to particle and multi-messenger astrophysics

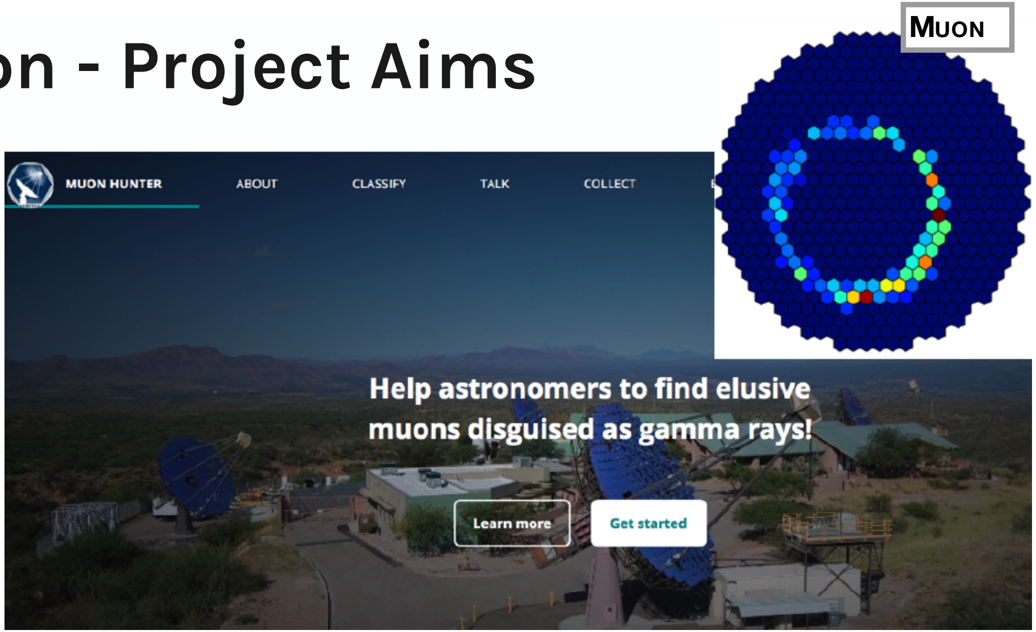
GWitch Hunters: correlating glitches with auxiliary channels to determine cause.

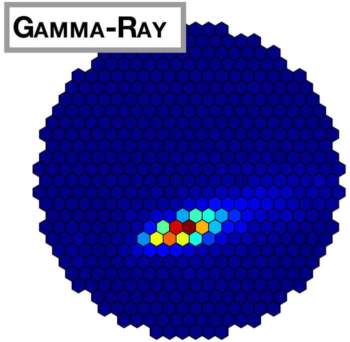


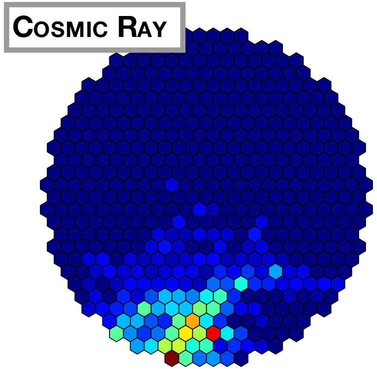
Muon tomography of volcanoes



Motivation - Project Aims





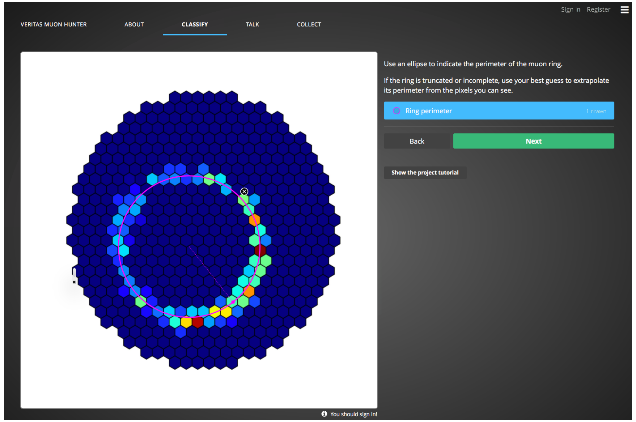


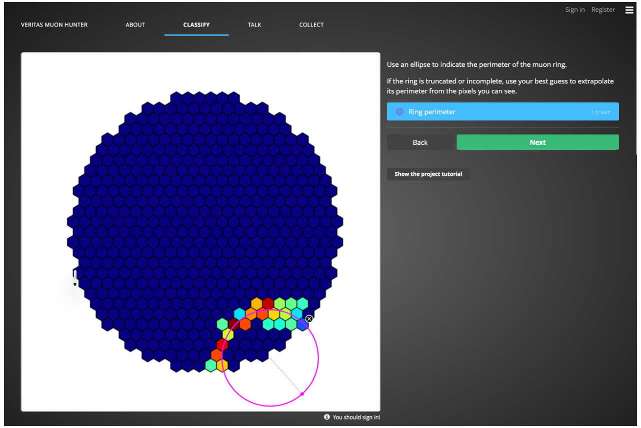
- **Ultimately**, to design and train a **machine classifier**** that can **distinguish all event classes**.
- **Initially**, focus on **muons** versus **non-muons**.
- Distinctiveness of **muon ring** images improves confidence of classifiers.

** **Convolutional Neural Networks** (CNNs) are well suited, but need abundant **labelled training data**.

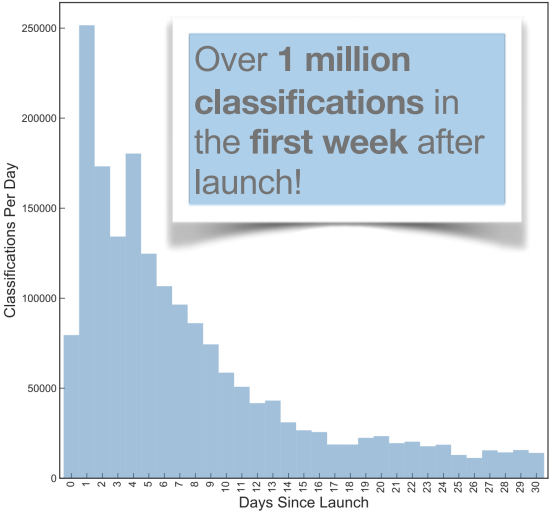
Muon Hunter

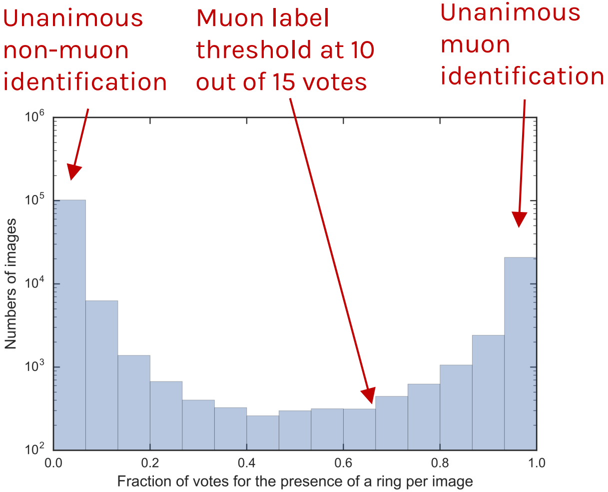
ZOONIVERSE





Fantastic public response!

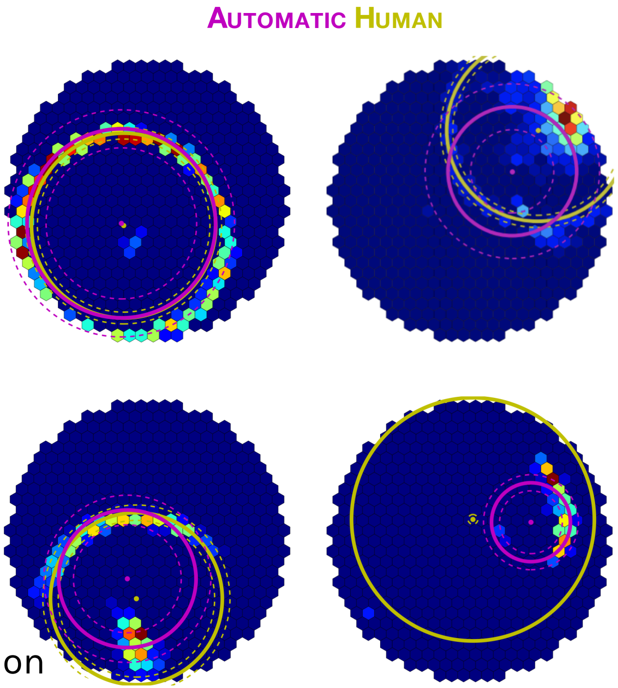




- **Completed** within **two months**.
- Final project statistics:
  - **2,161,338** classifications
  - **135,000** subjects classified
  - **6,107** volunteers.
  - **25,000** "muons"
  - **110,000** "non-muons"

Analysis: Consensus and Reliability

Good automatic performance is often matched or bettered by human classification.



Intrinsically faint muon rings worsen automatic results.

Human classification remains robust.

Additional Bright Pixels Impact Automatic Measurements.

Human Classification Remains Robust In Such Situations.

Outliers can impact human classification results…
BUT
Variance is a good indicator of outliers.

2. CNN training:

- Used labeled data to train simple VGG-based CNN
- Accuracies:
  - **95% - VEGAS labels**
  - **97% - MH labels** (67% agreement from humans)

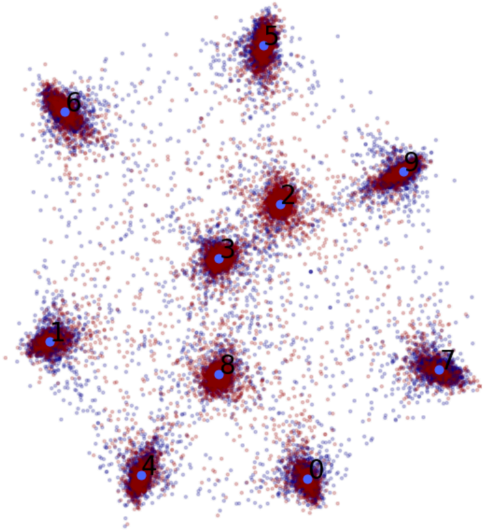
Q. Feng+ (2017), *A citizen-science approach to muon events in imaging atmospheric Cherenkov telescope data: the Muon Hunter*, 35th ICRC: https://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=301

R. Bird+ (2018), *Muon Hunter: a Zooniverse project*, TAUP 2017, DOI: 10.1088/1742-6596/1342/1/012103

1. Comparison of volunteer labels against standard VERITAS muon selection algorithm.

Take advantage of structure in unlabeled data

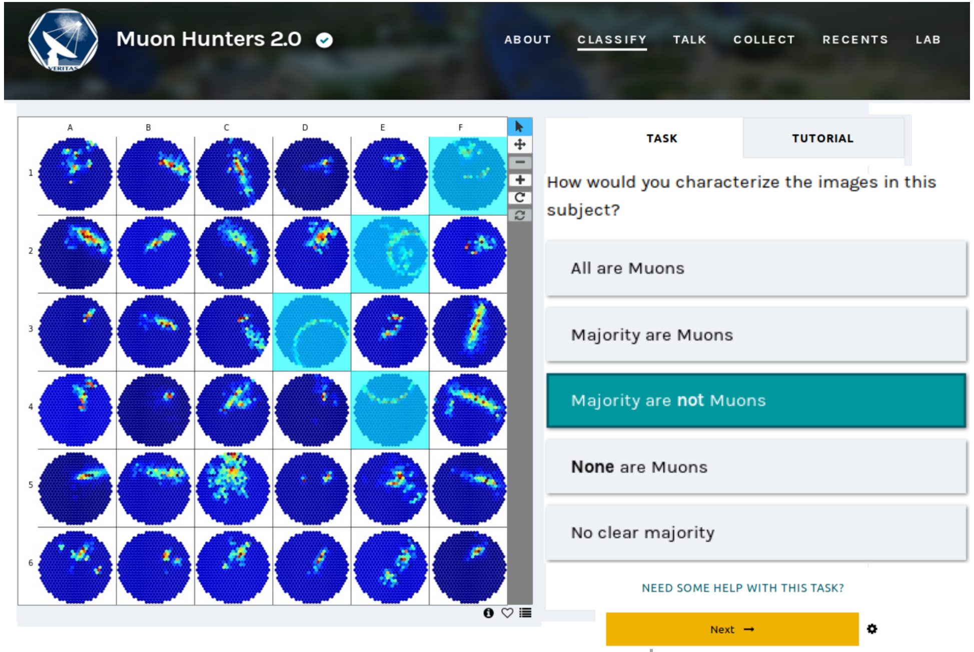
(a) DEC



Select subset of unlabeled cluster and pass to humans for labels

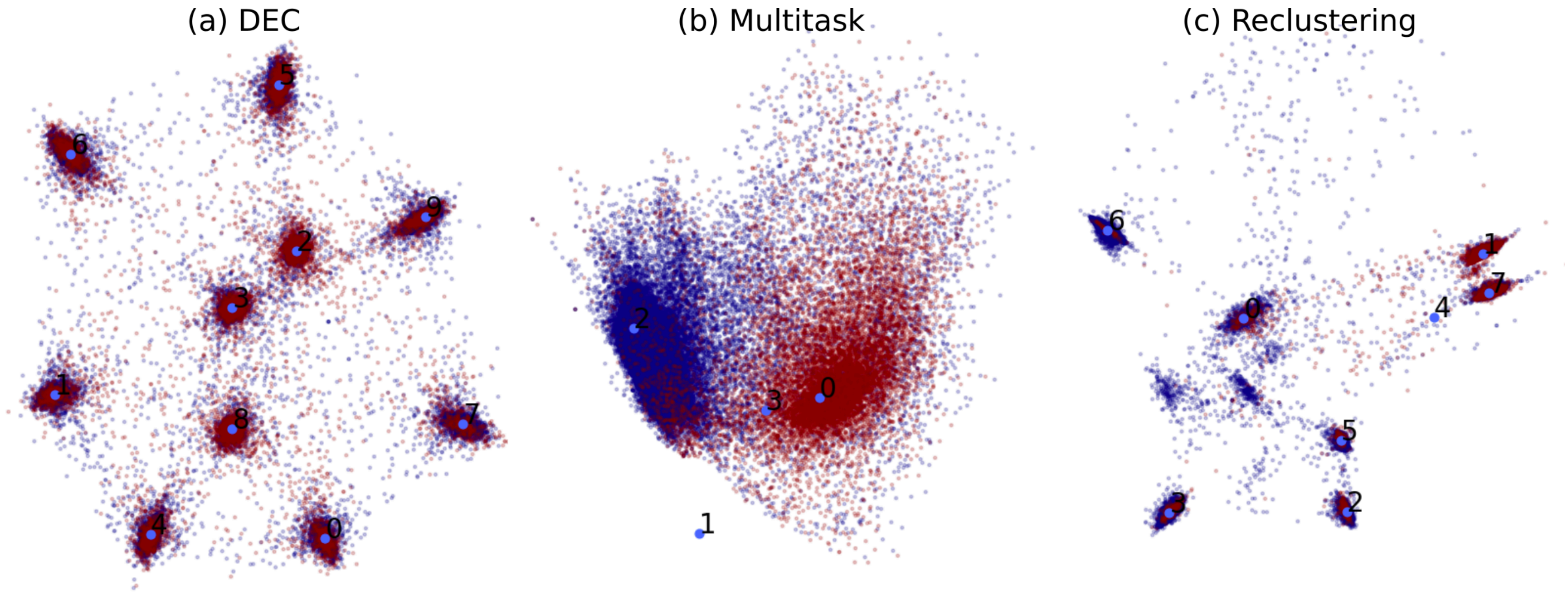
7.7 million labels obtained in two months.

20 X more efficient label gathering than single images!



- Use **unsupervised** machine learning to find **clusters of similar images**.
- Volunteers **classify** and **filter** clusters.
- Use volunteer annotations to **train** a **supervised** machine learning **model**.

Unsupervised clustering → multitask retraining → reclustering results in cluster purification which can then be resampled and labeled...

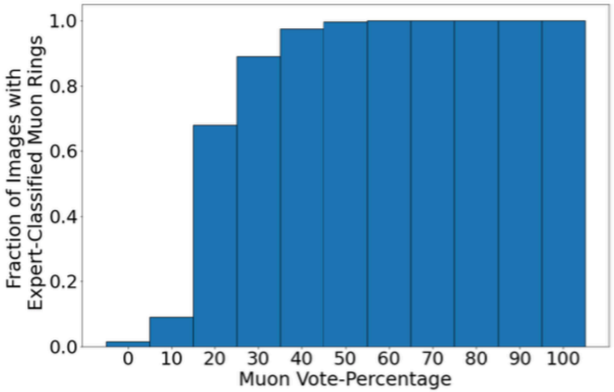


J. Xie, R. Girshick and A. Farhadi, *Unsupervised deep embedding for clustering analysis*, International conference on machine learning, pp. 478–487, **2016**.

D. Wright, M. Laraia, L. Fortson, C. Lintott and M. Walmsley, *Help me to help you: Machine augmented citizen science*, ACM Transactions on Social Computing, 2, **2019**

Laraia M., Wright D., Dickinson H., Simenstad A., Flanagan K., Serjeant S., Fortson L., et al.,, ICRC, 36, 678, **2019**

MH2 data can also be used again to benchmark simple muon detection CNNs – this time optimizing for the maximum number of pure muons.



Model	Accuracy (%)	Maximum number of pure muons
VEGAS	69.9	551 (11%) *(boundary = 0.024)*
MH2	95.7	3892 (78%) *(boundary = 0.064)*

Comparison of model performances on expert-labelled test set (5K each muon/not)

Identification method	Number of muon images identified
VEGAS algorithm	728
VEGAS-trained CNN	3071
MH2-trained CNN	23748

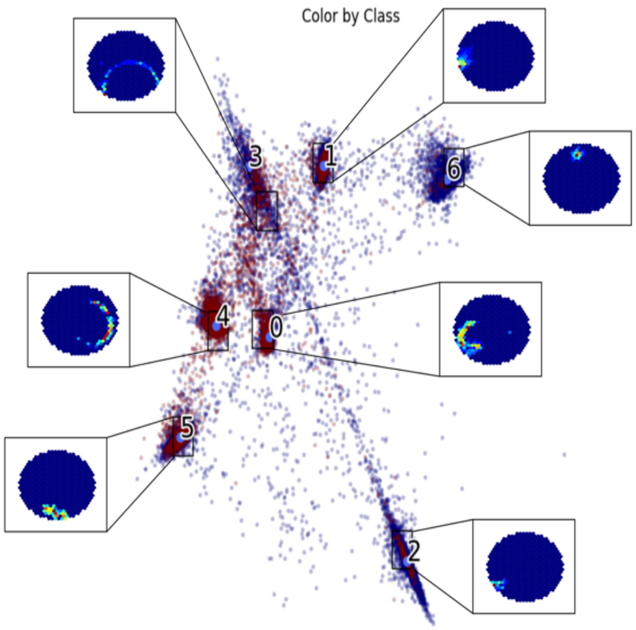
Number of muon images identified by each method out of the 481,819 images in the dataset.

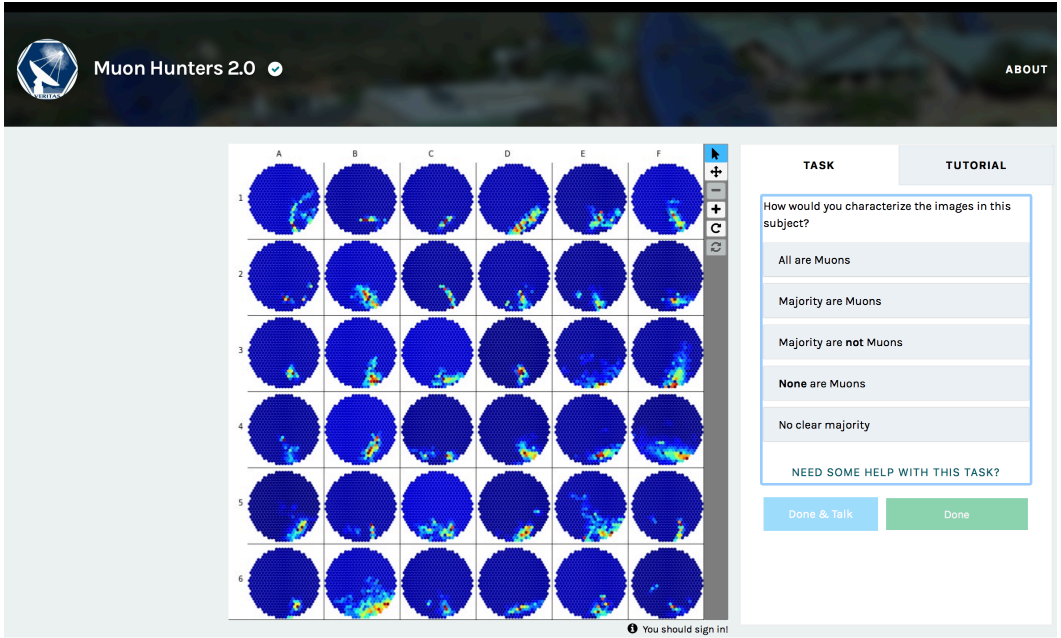
- Set vote threshold at 20% - removes most muons from “non-muon” class for model training
- Adjust class decision boundary for model until no false positive muons in validation
- MH2-trained model more accurate than VEGAS-trained model with much higher number pure muons.

In progress: implementing MH2-trained model in muon calibration pipeline.

K. Flanagan+ (2021), *Identifying muon rings in VERITAS data using convolutional neural networks trained on images classified with Muon Hunters 2*, 37th ICRC: https://arxiv.org/pdf/2108.07771.pdf

Unsupervised clustering algorithms along with anomaly detection can identify potentially novel classes which humans can then verify



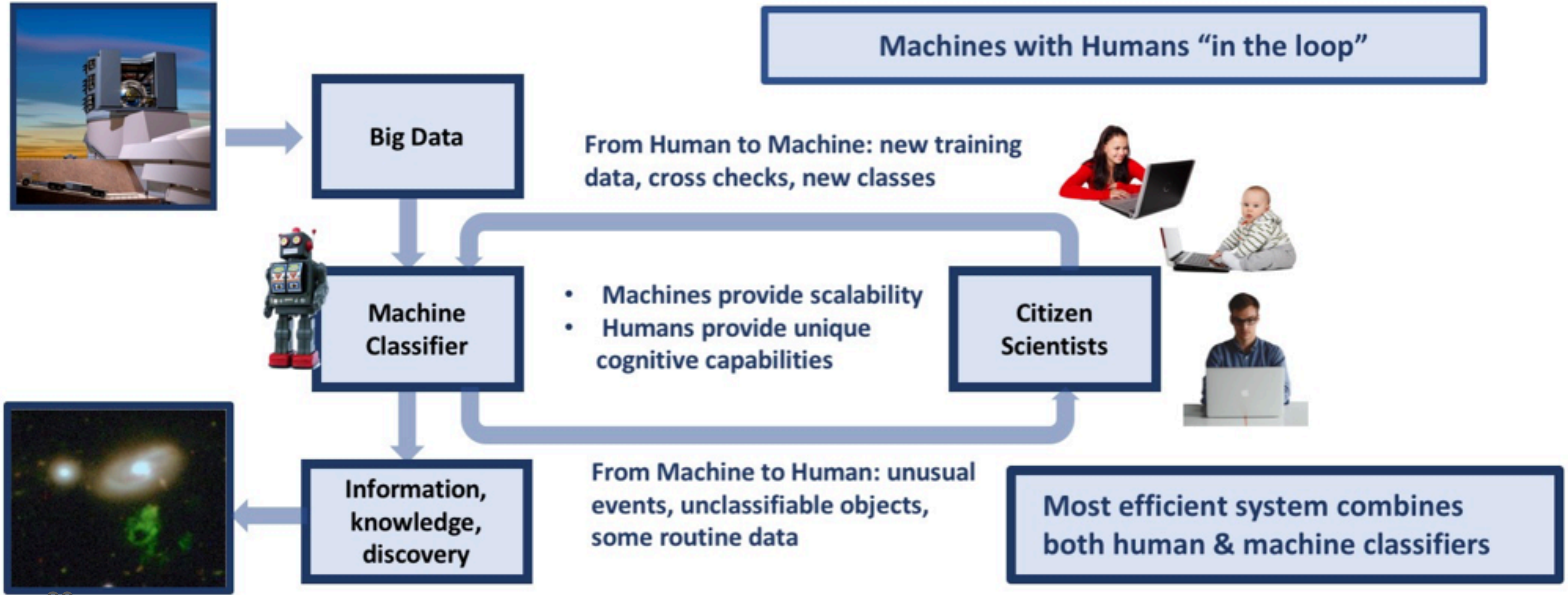


Anomaly detection work currently progressing with the Galaxy Zoo project – lessons learned will be applied to IACT data.

Zooniverse Capabilities Incorporating Machine Learning

- Track and use volunteer history to:
  - expedite retirement of subjects → **spacewarps.org**
  - send specific subjects to “expert” classifiers
  - “level-up” volunteers to harder workflows → **gravityspy.org**
- Track subject associated metadata to:
  - follow rules based on machine output → **galaxyzoo mobile**
- Some ML tested so far:
  - Object detection → **snapshot safari**
  - Active learning → **galaxyzoo**
  - Unsupervised clustering → **muonhunter**

Zooniverse is ready to help close the analysis gap in particle astrophysics and more!



And thanks to all of our volunteers!!!

Thank you!

lucy@zooniverse.org

@lucyfortson

LEVERHULME TRUST

NSF

NASA

ALFRED P. SLOAN FOUNDATION 1934

Google

Microsoft

SEVENTH FRAMEWORK PROGRAMME

Science & Technology Facilities Council

NOAA NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION U.S. DEPARTMENT OF COMMERCE

NATIONAL ENDOWMENT FOR THE Humanities

INSTITUTE of Museum and Library SERVICES

hhmi