# SPACE WARPS

ZOONIVERSE

## Crowd-sourcing the discovery of gravitational lenses

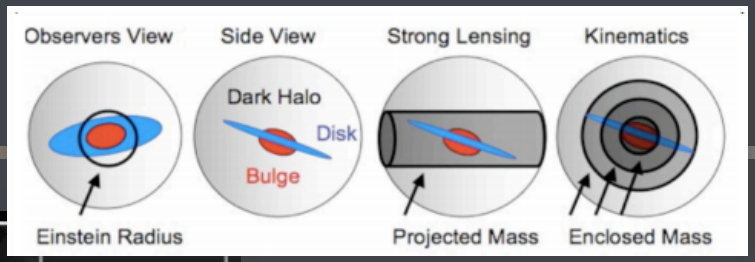
Kavli Symposium: Future Surveys and Citizen Science

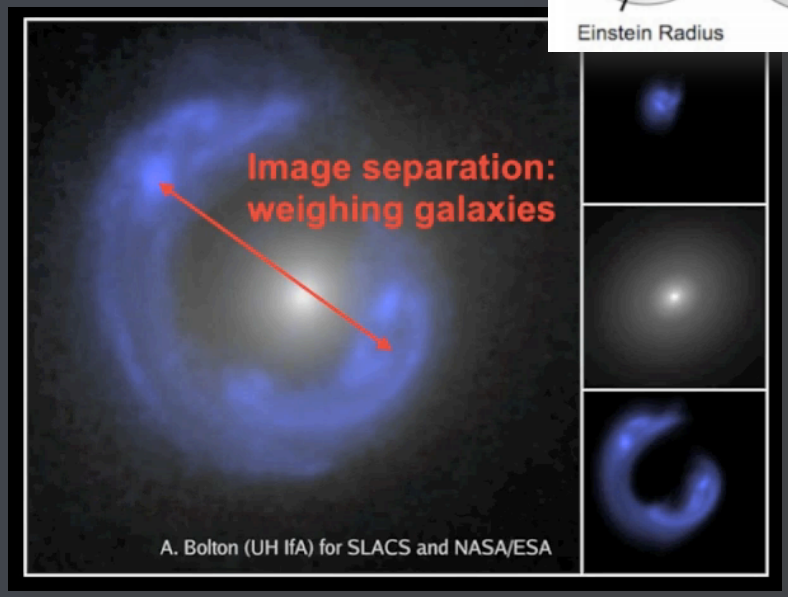
15th April 2015

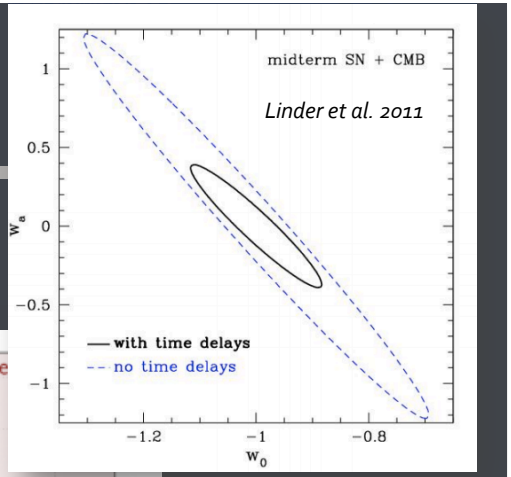
Aprajita Verma

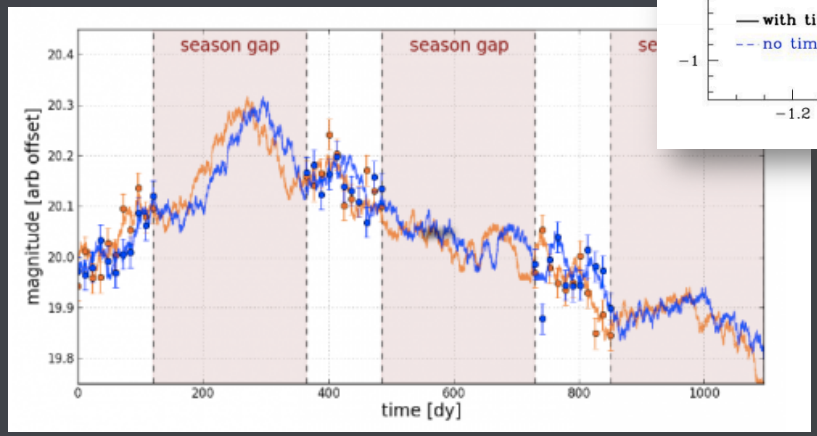
averma@astro.ox.ac.uk

## Weighing galaxies



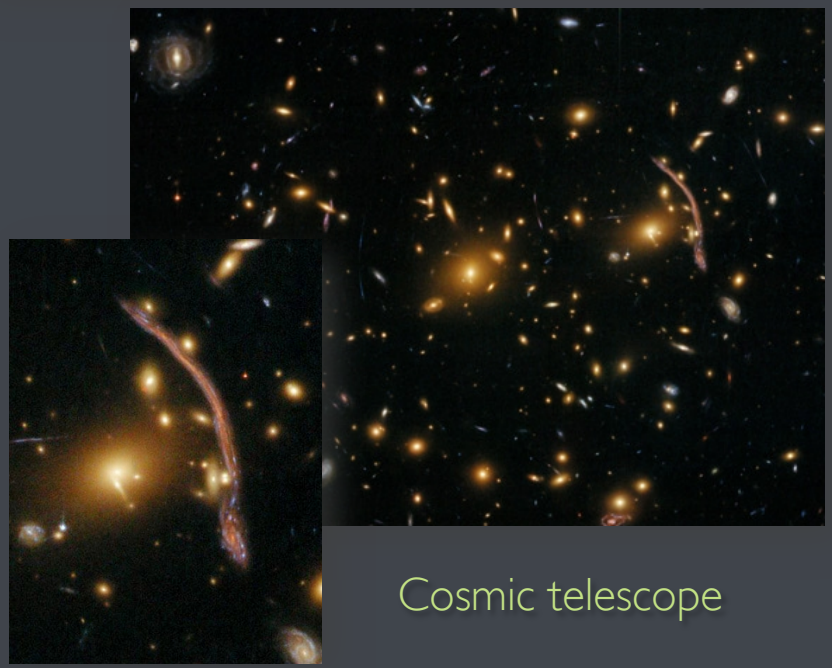



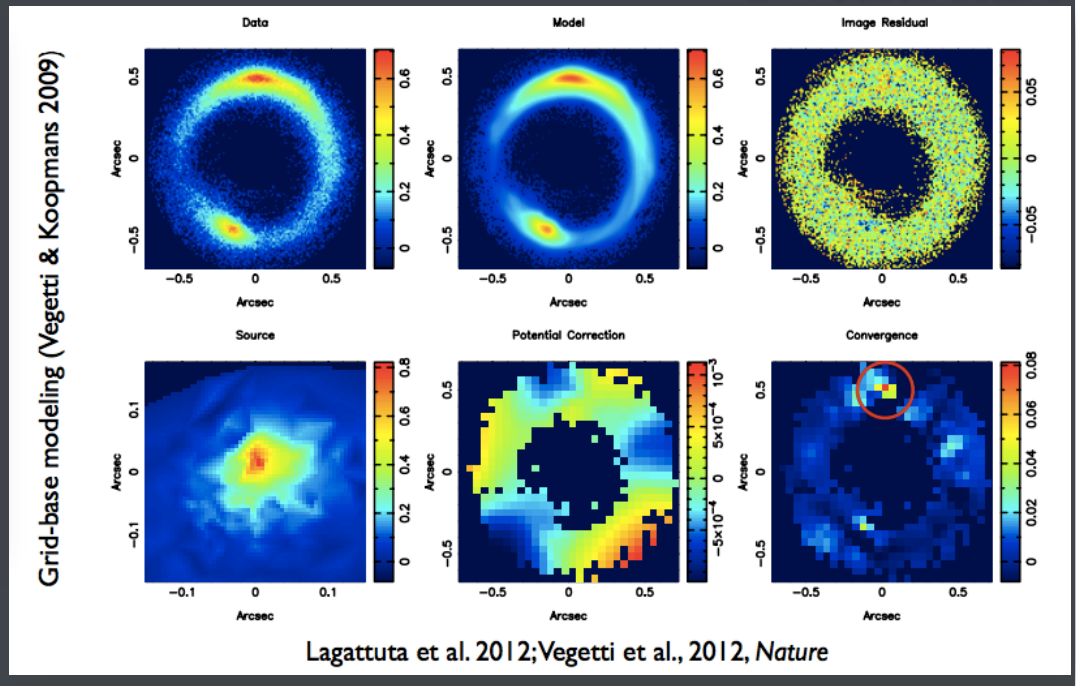




## QSO/SN time delays

## DM Substructure



## Cosmic telescope




Lagattuta et al. 2012; Vegetti et al., 2012, *Nature*

# Needles in a Haystack

## Strong lensing is rare

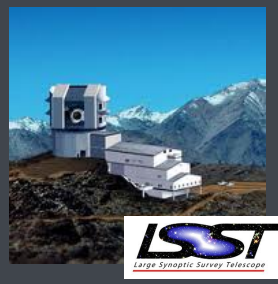
only 1 in ~10000 massive galaxies act as lenses

only ~ 500 strong lenses currently confirmed

# Needles in a Haystack

entering an era of very large & sensitive imaging surveys

~8000 lensed AGN, $10^{4\text{-}5}$ galaxy-scale lenses
1000s of clusters, 100s of lensed SNe...




# How will we find them?

Automated lens finding algorithms (robots) have been developed with good success


But robots might not find everything ….

- good at finding types of lenses (e.g. "ring-finder" "arc finder")
- may miss low surface brightness or low separation images
- complex or "exotic" lenses

Even with "robots" in place, strong gravitational lensing research remains very labour-intensive:

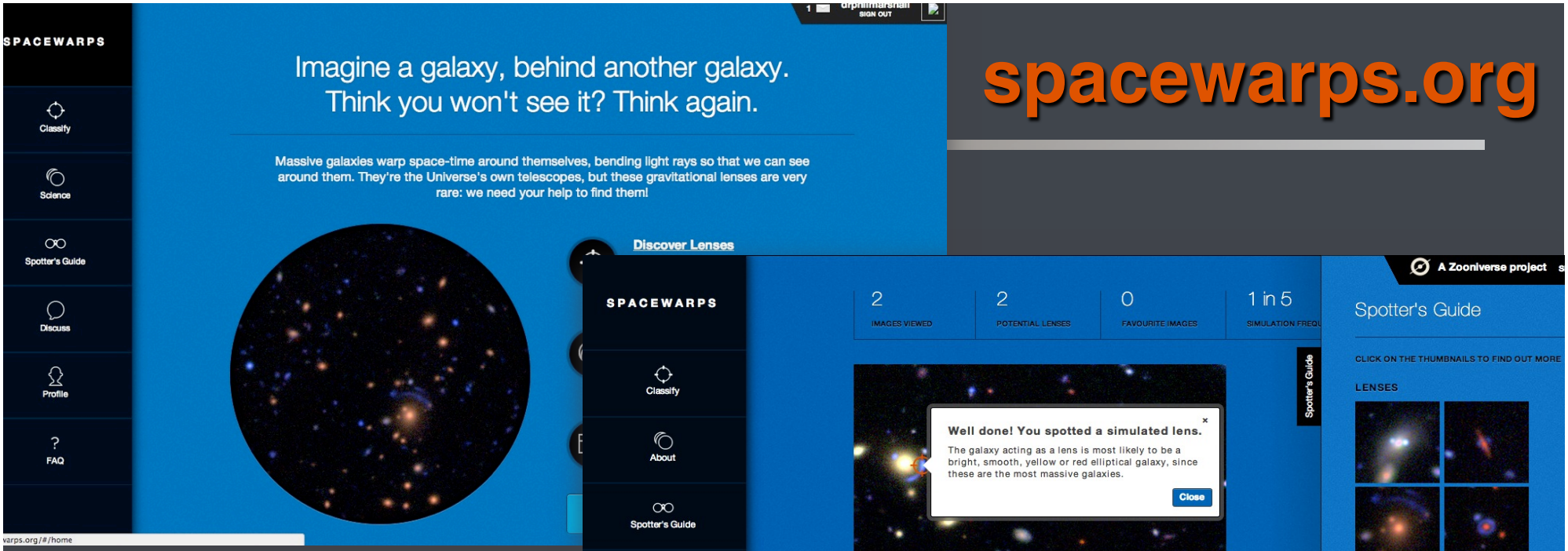
- Inspection of lens candidates from robots
- Detailed modelling of individual systems

From http://www.vanderstahl.com

# The Solution: Part 1

http://commons.wikimedia.org


EN
1,112,192 people taking part worldwide
Projects
Community
About us
Sign in
Register
ZOONIVERSE
We make citizen science websites so that everyone can be part of real research online
THE SOLUTION: PART 2
condorwatch.org
Help save the California condor!
All
Space
Climate
Humanities
Nature
Biology
www.zooniverse.org
Space
Sort by
Category
How do galaxies form?
NASA's Hubble Space Telescope archive provides hundreds of thousands of galaxy images.
Explore the surface of the Moon
We hope to study the lunar surface in unprecedented detail.
Study explosions on the Sun
Explore interactive diagrams to learn about the Sun and the spacecraft monitoring it.
Find planets around stars
Lightcurve changes from the Kepler spacecraft can indicate transiting planets.




# spacewarps.org

- SW1: CFHTLS: blind search
  - 430,000 images, (p0 ~ 1/5000)
  - Understand selections function
  - Improve automated algorithms
  - Rare objects: lensed quasars, unusual/exotic (faint, red, odd-shaped…)
    - arcs in optimized color composite images
  - Training set of "sims" and "duds" used to teach volunteers what to look for, and what to ignore

# SPACE WARPS Collaboration


## Space Warps Science Team:

PIs: Phil Marshall (KIPAC Stanford) Anupreeta More (IPMU) Aprajita Verma (Oxford)

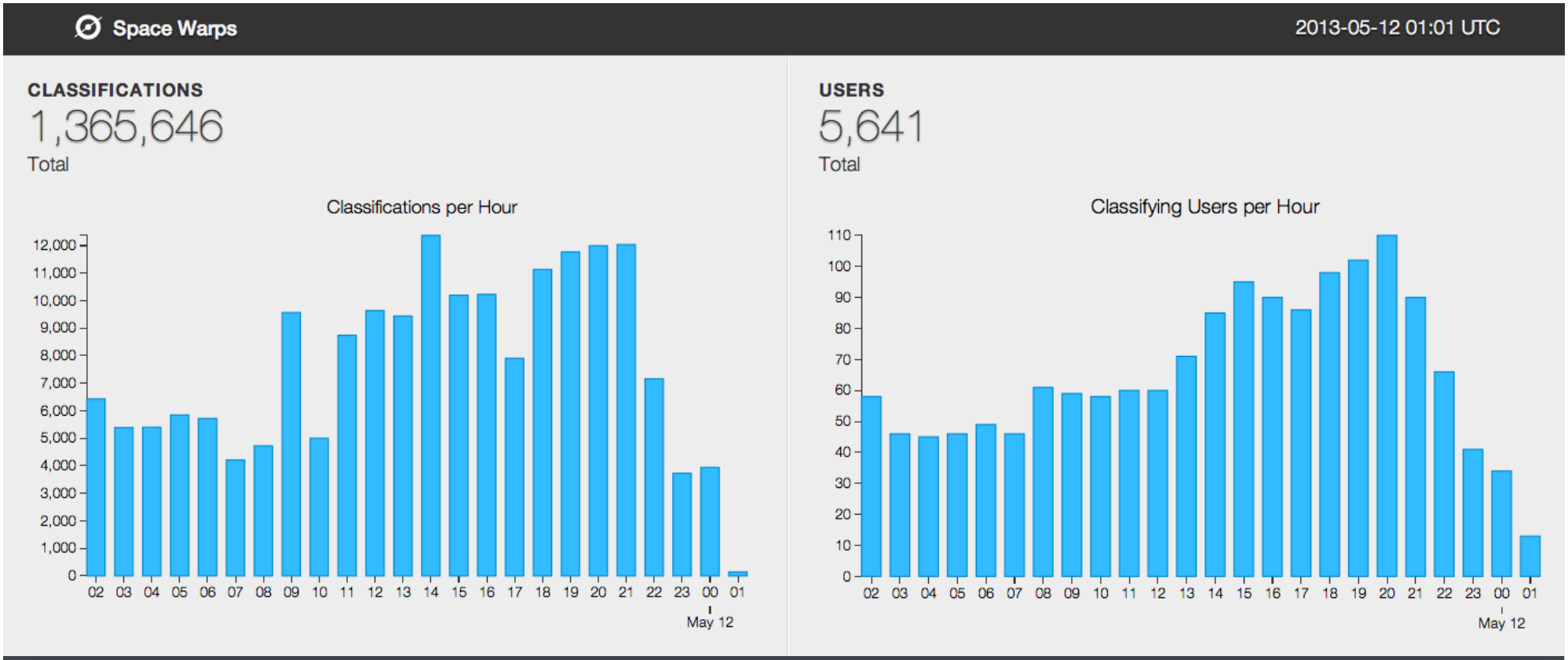
Elisabeth Baeten,* Claude Cornen,* Thomas Jennings,* Rafi Kueng (Zurich), Christine MacMillan,* Surhud More (IPMU), Julianne Wilcox,* Layne Wright,* Prasenjit Saha (Zurich), Matthias Tecza (Oxford)

## Space Warps Development Team:

Amit Kapadia, **Chris Snyder, Michael Parrish**, David Miller, Arfon Smith, Kelly Borden (Zooniverse/Adler Planetarium), Rob Simpson, Chris Lintott (Oxford)

## VICS82 Survey Team:

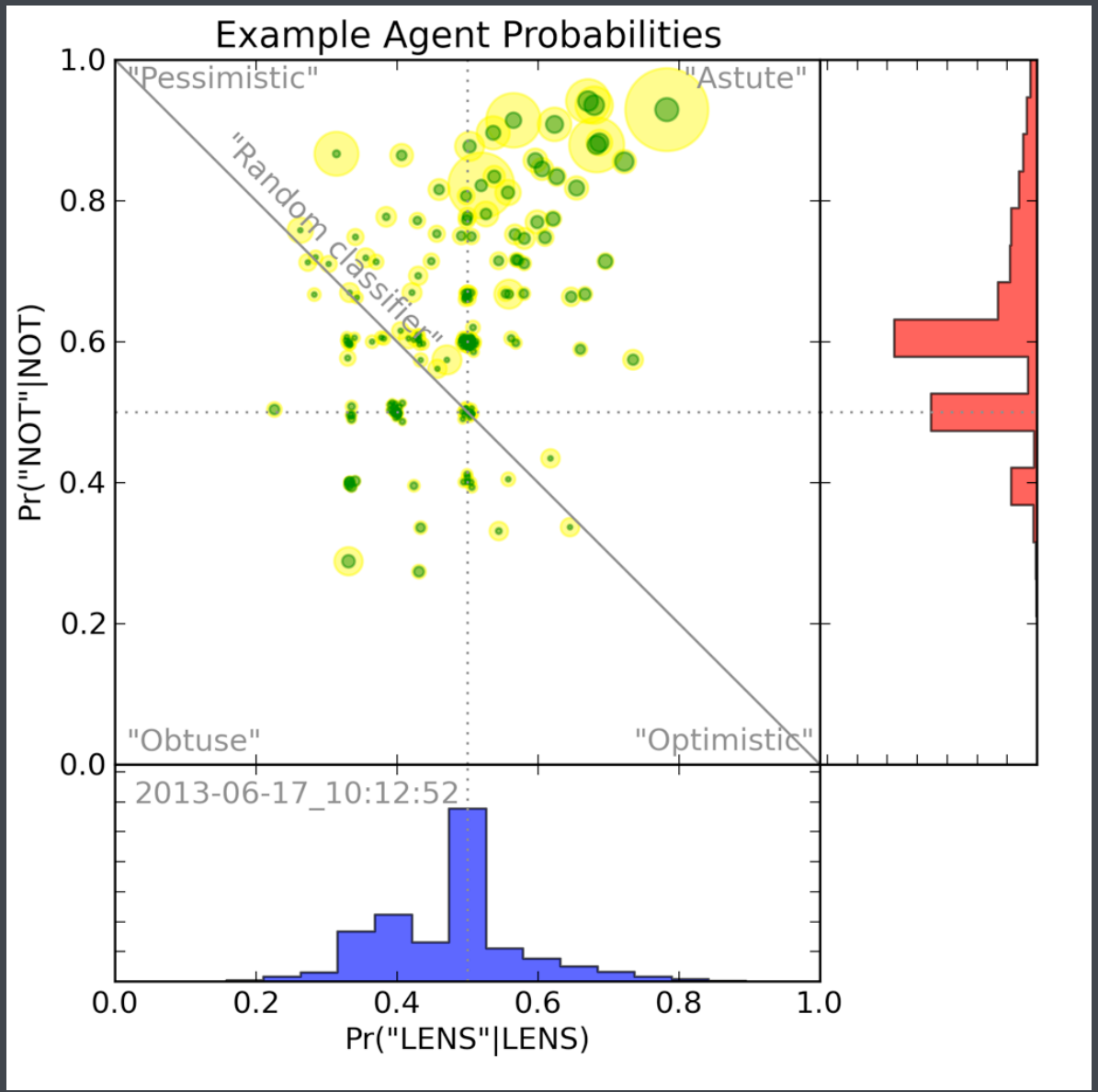
Jim Geach (Herts), Jean-Paul Kneib (EPFL), Yen-Ting Lin, Martin Makler (CBPF), Eli Rykoff & Eduardo Rozo (KIPAC)




- 2 million image classifications in first week (10k per hour), 11 million over 6 months
- Collaboration of over 60,000 citizen scientists, but typically only ~10-100 active at a time

# SWAP: Making sense of the markers

- Assign each volunteer a software Agent, which: listens to its volunteer classifying images as "LENS" or "NOT" by marking (or not)
- Each Agent tries to describe its volunteer with just two numbers:
  - Pr("LENS"|LENS)
  - Pr("NOT"|NOT)
- Agent updated each time a training image is classified
- Encodes all our assumptions; Humans are complex, Agents are simple
- Errors: mistake classifications, our interpretation

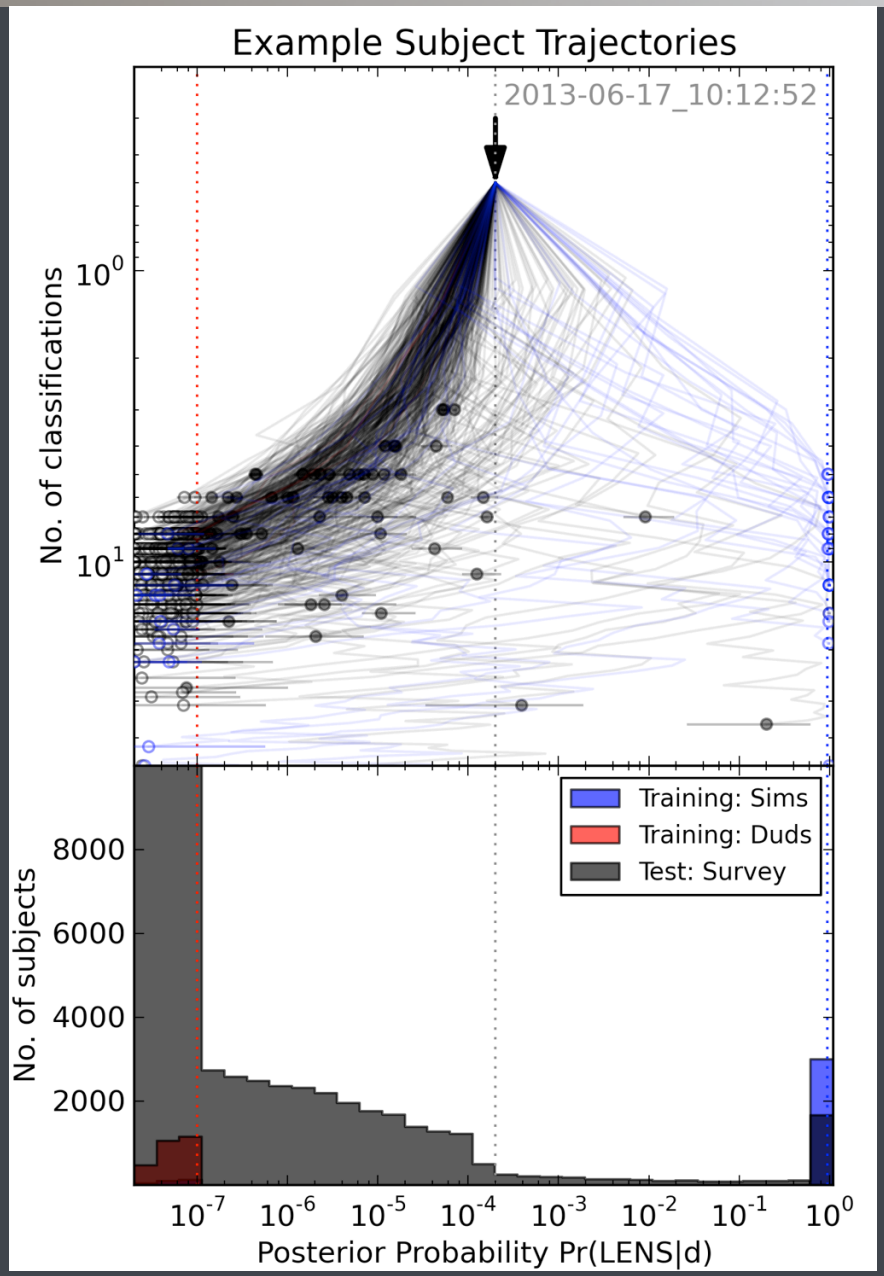


# SWAP: Making sense of the markers

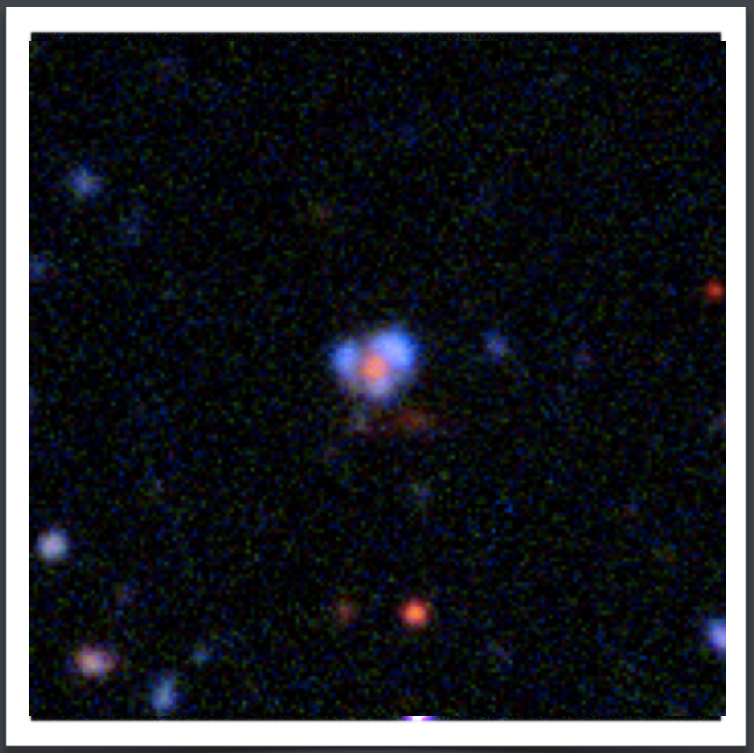
- Agents update the posterior probability
- Pr(LENS|d) for each survey image ("subject") as it is seen, in an online inference

$$\Pr(\mathrm{LENS}|C_k, T_k) = \frac{\Pr(C_k|\mathrm{LENS}, T_k) \cdot \Pr(\mathrm{LENS})}{[\Pr(C_k|\mathrm{LENS}, T_k) \cdot \Pr(\mathrm{LENS}) + \Pr(C_k|\mathrm{NOT}, T_k) \cdot \Pr(\mathrm{NOT})]} \quad \text{(A4)}$$

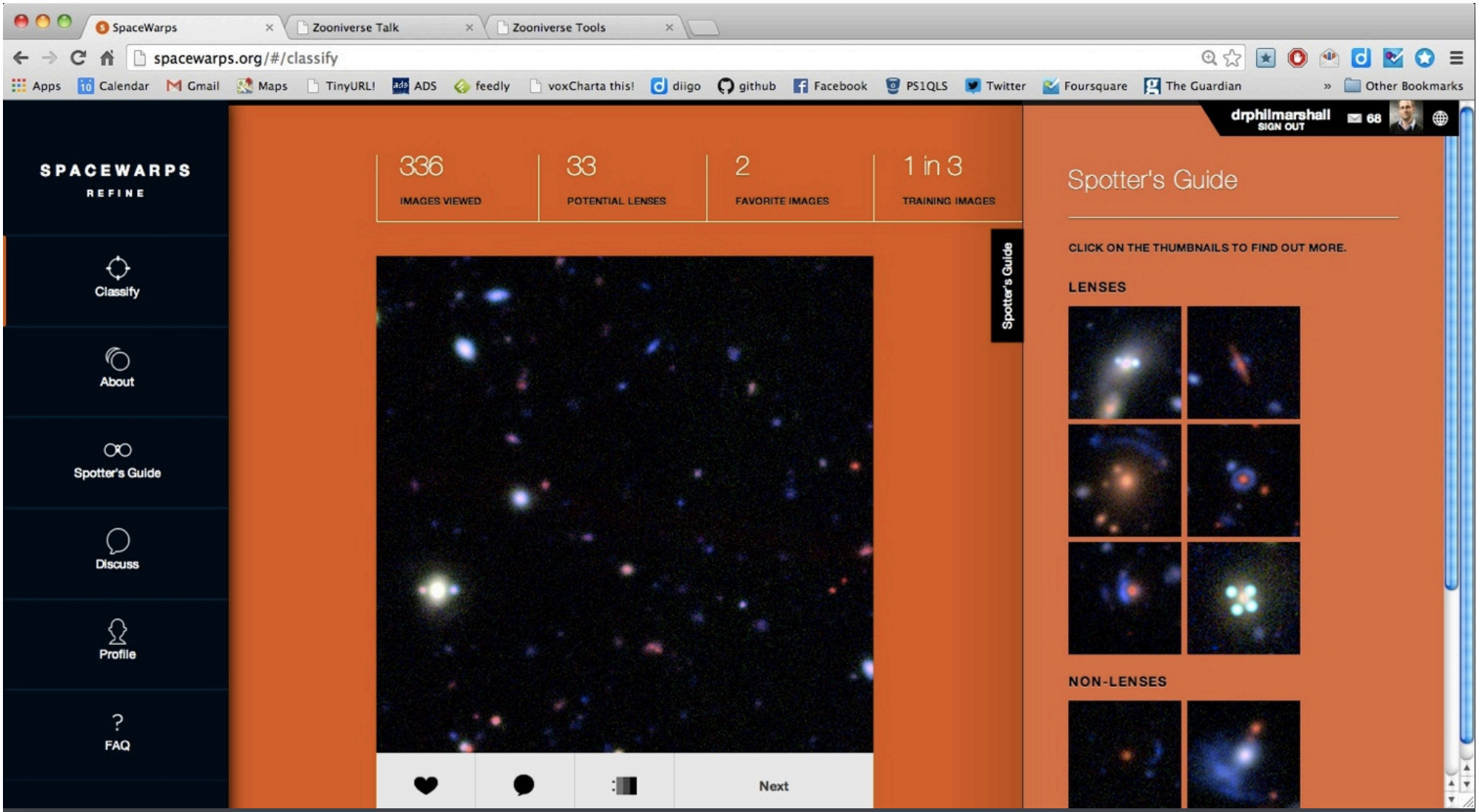
- Below p=$10^{-7}$, subjects are retired from the stream
- Above p=0.95, subjects are labelled "candidates"
- ~10 classifications needed per rejection



# CFHT-LS Stage I Results

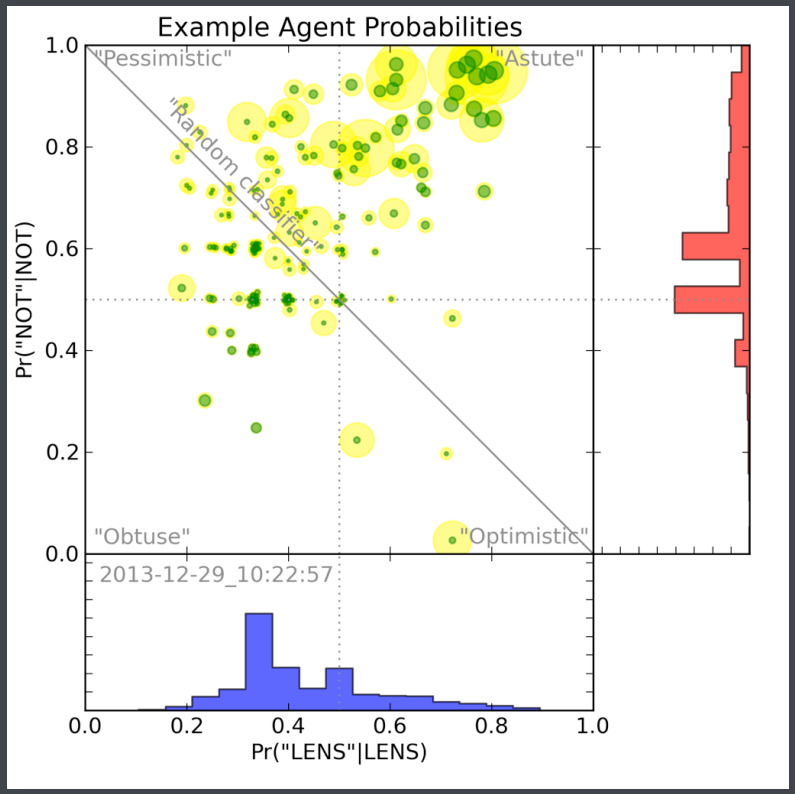


- Completeness (to sims) is 93%, with missing 7% understandably non-visible
- Some good candidates identified: galaxy-scale systems that would fail an LRG cut, faint group-scale arcs
- ~3381 candidates, out of 430,000 images (99.2% rejections)

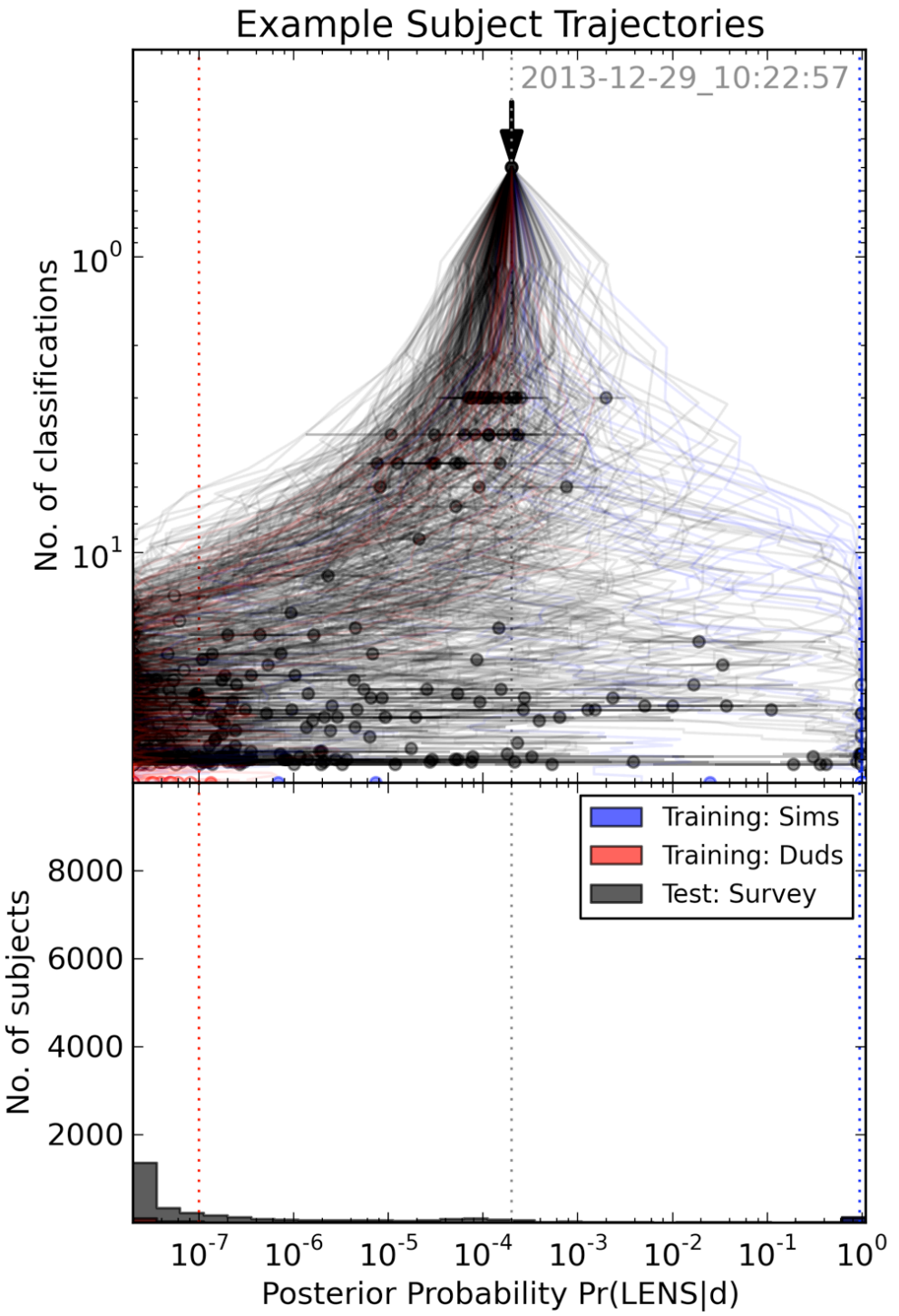


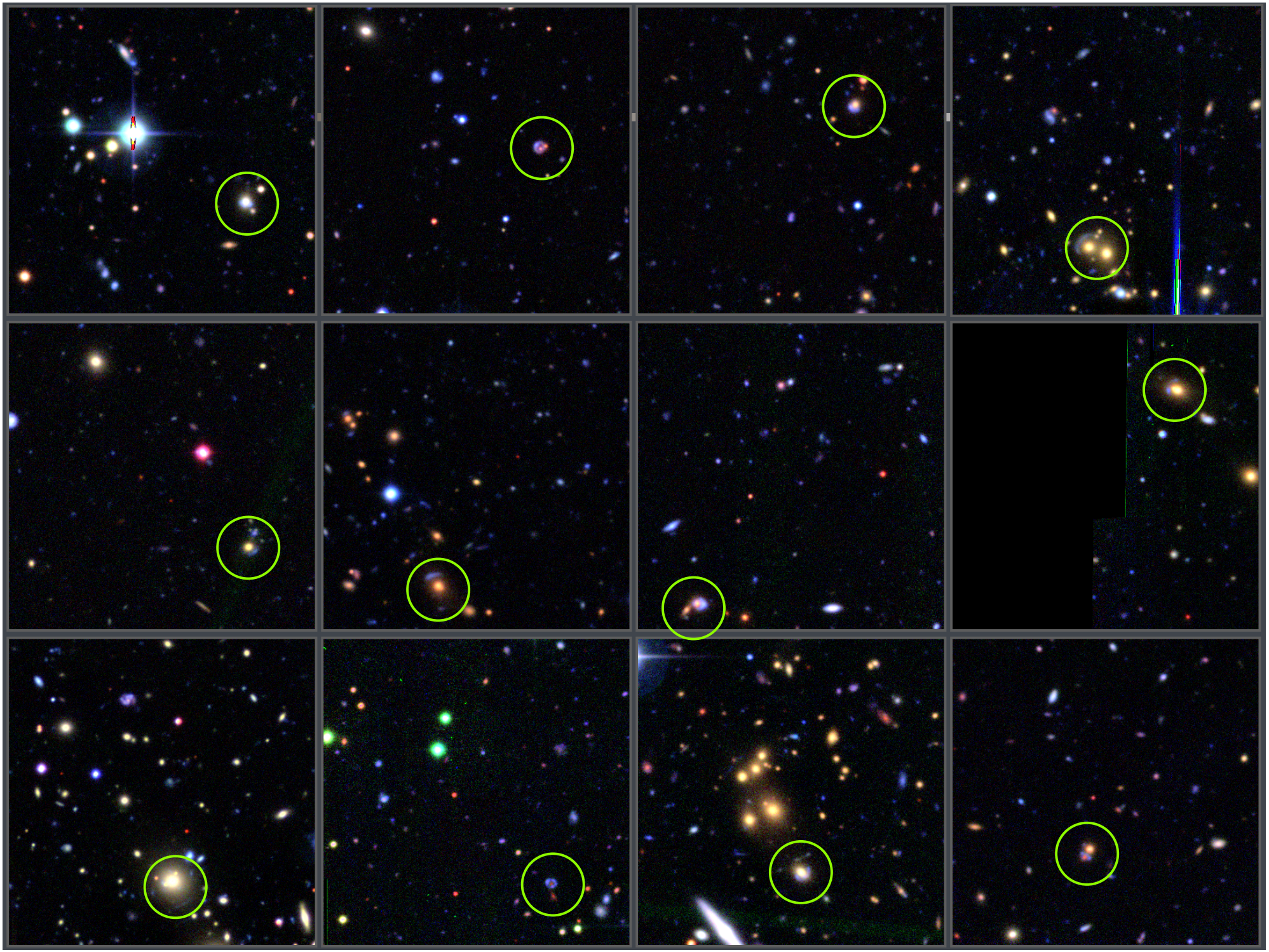

- Stage 2 sample "refinement:" science team reconfigured website (via github), painted it orange
- New "harder" training set with feedback, sims & false positives

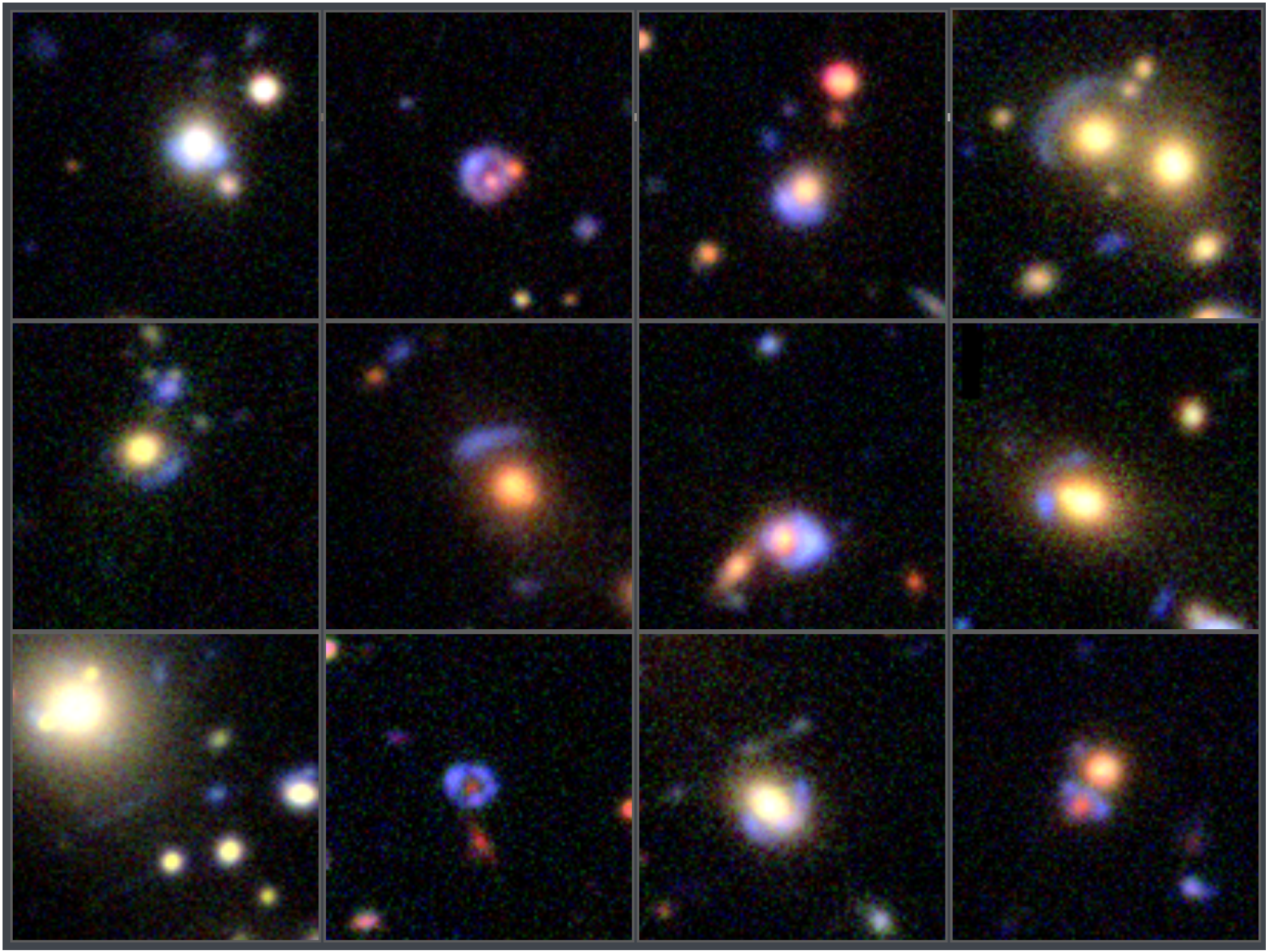
# Stage 2 First Results



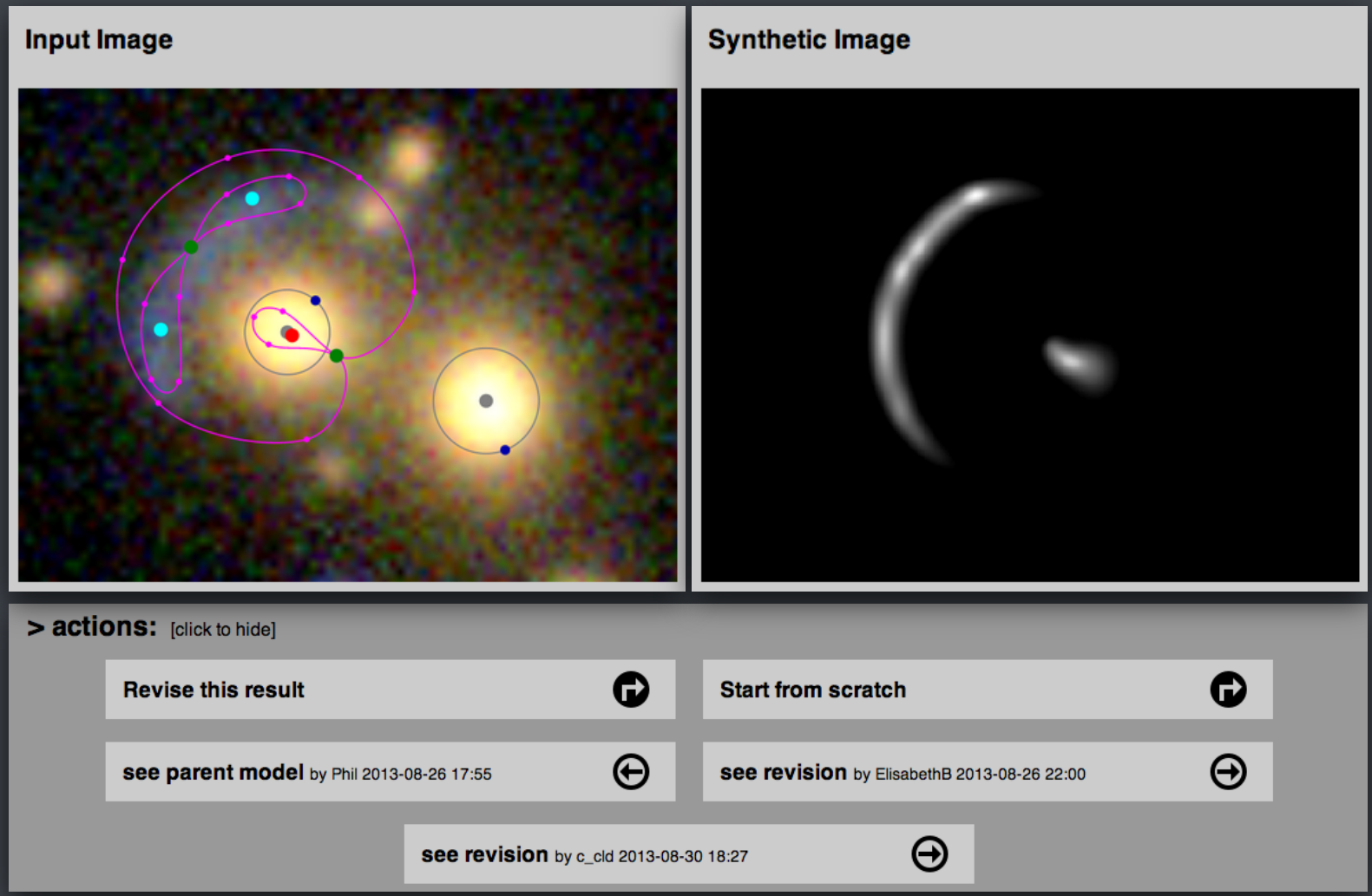

- Ranked list of candidates by P
- Completeness (in harder sims) is 78%
- "Expert" inspection of top 600
- ~60 new lens candidates





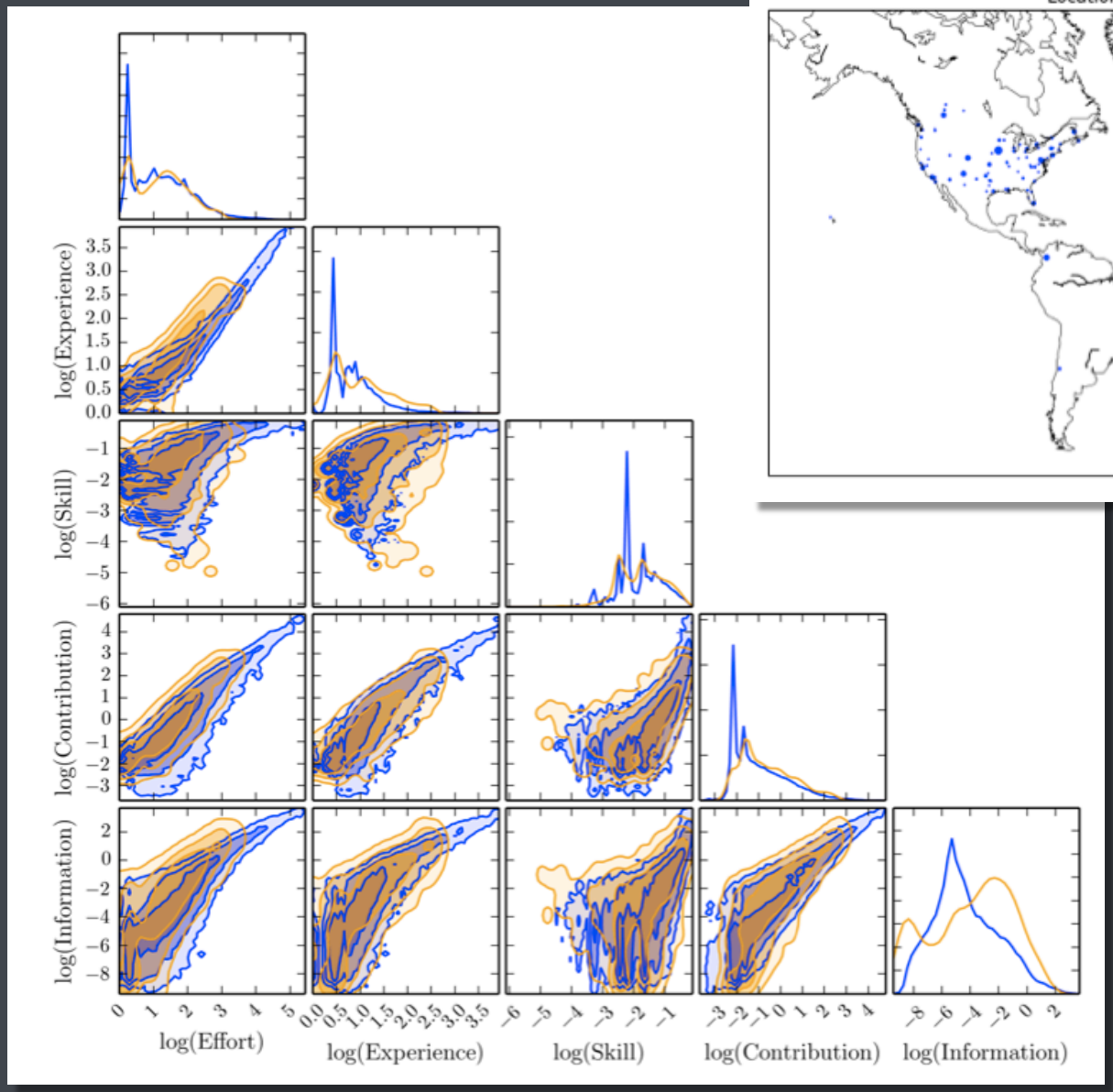


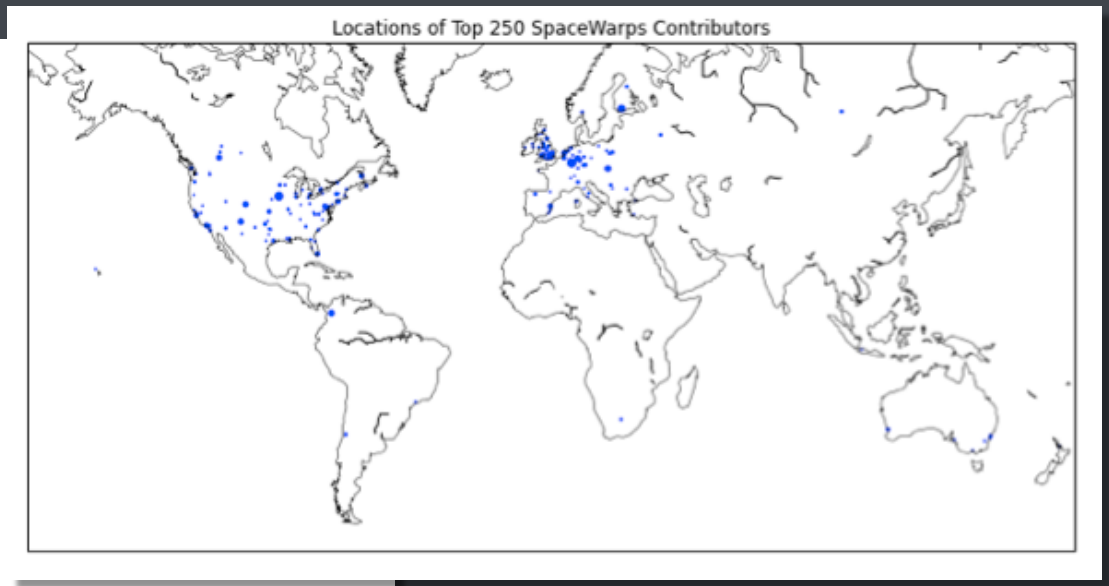
# Citizen Lens Modelling

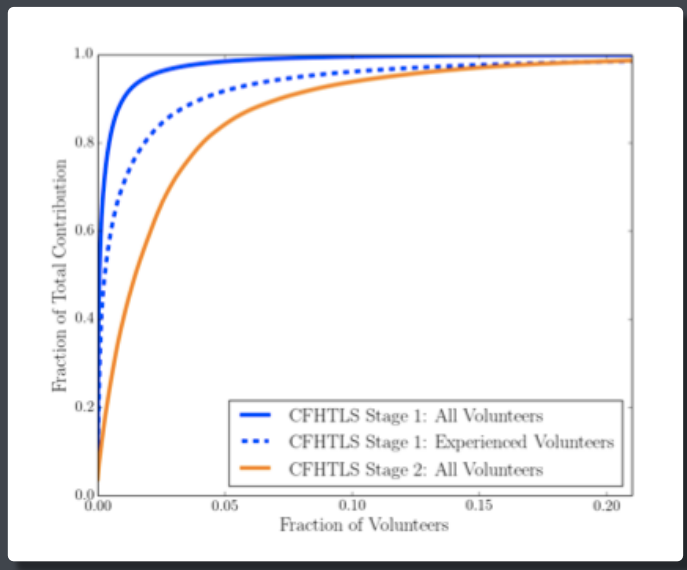



"SpaghettiLens"
Kueng et al. MNRAS submitted

# Understanding the crowd



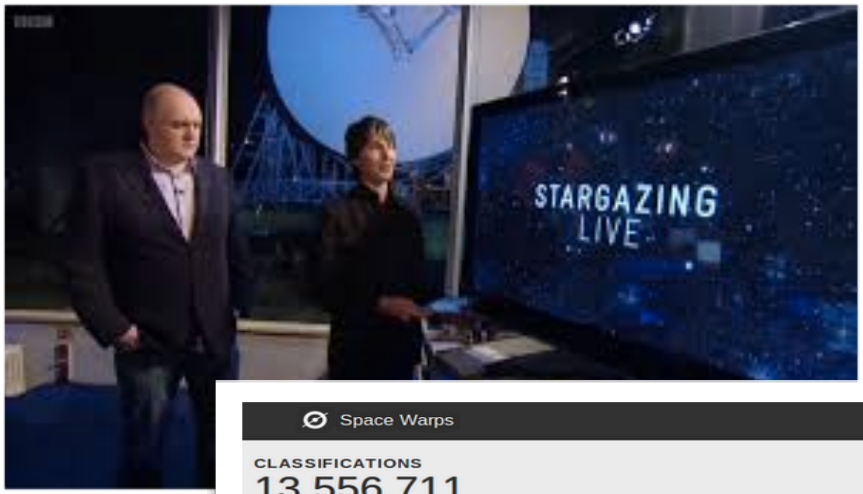



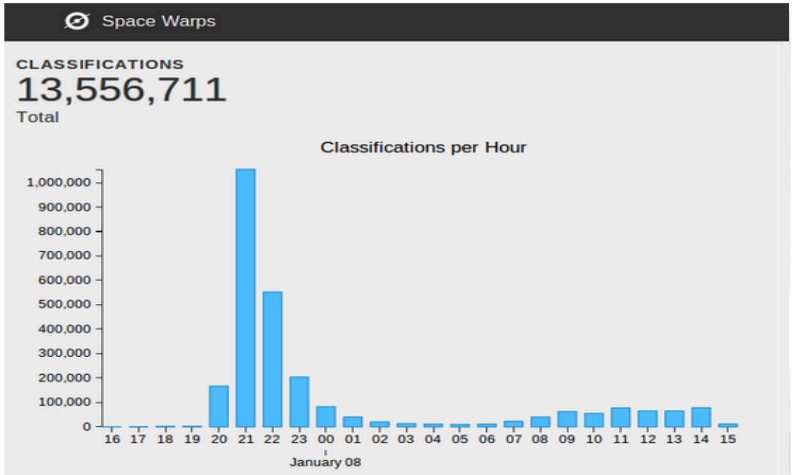

*Chris Davis, Michael Baumer, Phil Marshall*

# SPACE WARPS 2: VICS82

40,000 “targets” selected from SDSS++

iJKs images, sims generated as for CFHTLS, color scales chosen for maximum contrast

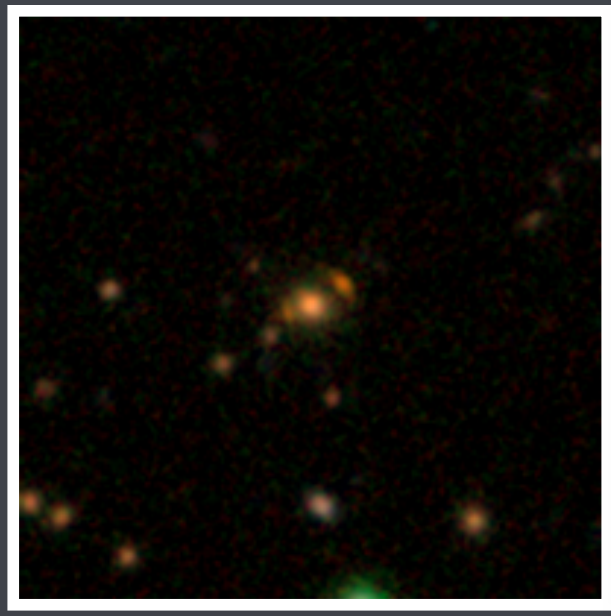


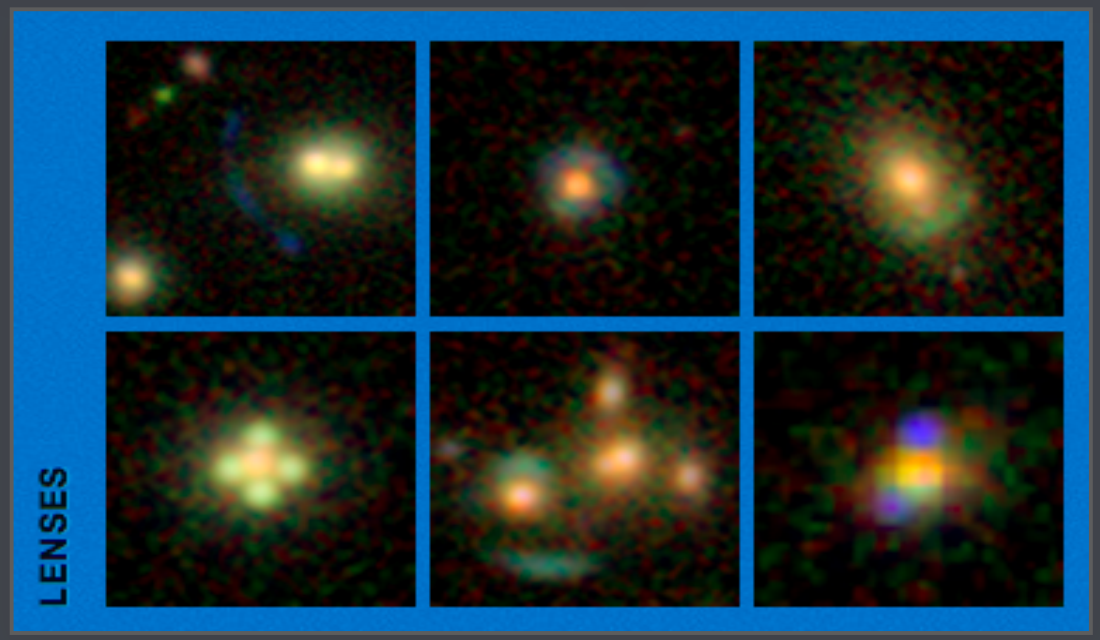




Launched on BBC Stargazing Live TV programs

1 million image classifications in the single hour after the first Stargazing Live show finished

# SPACE WARPS 2: Initial Results






- The 3 candidates have red arcs - despite the training sims only having blue/green arcs
- Space Warps is more than just a supervised learning system ... classifiers have imagination and intuition

# SPACE WARPS 1 & 2: Conclusions & Future

- Crowd-sourced lens detection successful (targeted and blind)
  - 90% complete (to sims) , 99% rejection
  - Find exotic or tricky lens candidates missed by robots
- Crowd are imaginative, intuitive and adaptable
- Crowd like to 'Talk' and to 'Learn', motivated by contribution to science
- Media fuelled classification frenzy (SW2): $10^6$ classifications per hour – LSST 20,000 sq deg in 40 days! (cf SW1 max 10k/h 1 1y)
  - How to maintain momentum? New unseen data, LSST-TV support language and time zones
  - Focus on sub-samples: large-scale QC inspection of robot candidates (targeted)

# SPACE WARPS I & 2: Conclusions & Future

- Completed first system & results papers
  - (Marshall, AV, More et al, More, AV, Marshall et al., Geach, More, AV et al.)
  - comparisons: humans vs robots, supervised vs unsupervised learning
  - analysing the Bureau: Where is the information coming from?

- COMING SOON: watch out for SPACE WARPS 3 in the autumn & more….

- SW3+: Interest from KiDS, DES (SW3), PS1, HSC, (VIDEO) …
- Space Warps Labs to port modeling to the crowd
- Euclid & LSST …