

WHEN SCIENTISTS BECOME SOCIAL SCIENTISTS: HOW CITIZEN SCIENCE PROJECTS LEARN ABOUT VOLUNTEERS

PETER T. DARCH

SCHOOL OF INFORMATION SCIENCES, UNIVERSITY OF ILLINOIS AT
URBANA-CHAMPAIGN

12TH INTERNATIONAL DIGITAL CURATION CONFERENCE
EDINBURGH, UK, FEBRUARY 21ST 2017



ILLINOIS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Citizen Science: Data Quality Challenges

- Citizen science projects:
 - Help scientists meet data curation challenges
 - Engage members of the public in the scientific process
- Citizen science is growing
 - *Zooniverse* comprises 52 projects, > 1 million volunteers
- **Projects are very concerned about data quality**



http://www.discoverthebluedot.com/gfx/profiles_239.png

Data Quality and Volunteer Management

- Volunteer management is vital for data quality
 - Volunteer recruitment
 - Volunteer retention
 - Volunteer behaviour
 - Volunteer abilities
- Volunteer management is shaped by a project's features



<http://astrobites.org/wp-content/uploads/2012/01/zoo1.png>

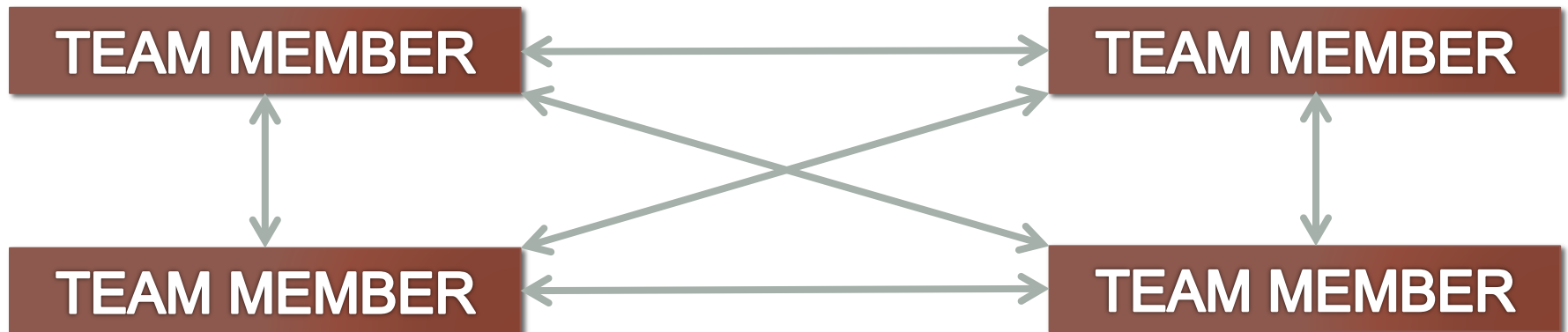
Importance of Learning About Volunteers

- Projects make decisions about what features to include
- Effective decisions require learning about volunteers
- Challenges of learning about volunteers
 - How to gain new insights about volunteers



Importance of Learning About Volunteers

- Projects make decisions about what features to include
- Effective decisions require learning about volunteers
- Challenges of learning about volunteers
 - How to gain new insights about volunteers
 - How to circulate insights among project team members



Research Questions

- 1. What do teams who operate citizen science projects learn about their projects' volunteers?**
- 2. How do these teams learn about volunteers?**
- 3. How can these learning processes be improved?**

Introducing Galaxy Zoo

- Launched July 2007
- Volunteers classified shapes of one million galaxies
- Over 1,000 papers cite Galaxy Zoo datasets
- Social scientific case study:
 - Interviews, observation, document analysis
 - Focuses on first four years of Galaxy Zoo

galaxyzoo.org



Learning About Volunteers: Three Moments of Learning

1. Credit system and volunteer motivations
2. Serendipitous discoveries and the *Green Peas*
3. Simulated images and bias testing



[http://zoo1.galaxyzoo.org/poster/
GalaxyZoo_lowres.jpg](http://zoo1.galaxyzoo.org/poster/GalaxyZoo_lowres.jpg)

Learning About Volunteers:

1. Credit Systems and Volunteer Motivations

- Galaxy Zoo launched with a league table
- Team members became aware that methods for crediting contributions:
 - Affect volunteer behaviour
 - Affect volunteers in different ways
 - Have implications for data quality

Recent Galaxy Classification Leaders

User Name	Classifications
Joseph K. H. Cheng	385115
CARL R.E. MILLS	382857
Hennex	267861
ALAN MASON	225194
ntel25	200800
dougre	169204
JEFF LAUER	163426
didi0815	139456
glanserse	137176
ElisabethB	127627
Magic	126031
clarea13	109501
Lytala	96429
b.umberto	90480
olddan	85072

Learning About Volunteers:

1. Credit Systems and Volunteer Motivations

- Team identified three groups of volunteers:
 1. Those motivated to improve table position
 2. Those demoralised by the league table
 3. Those unbothered by the league table

Learning About Volunteers:

1. Credit Systems and Volunteer Motivations

- Team identified three groups of volunteers:

1. Those motivated to improve table position
2. Those demoralised by the league table
3. Those unbothered by the league table

Website metrics

Email from volunteers

Forum posts



Learning About Volunteers:

1. Credit Systems and Volunteer Motivations

- Team identified three groups of volunteers:
 1. Those motivated to improve table position
 2. Those demoralised by the league table
 3. Those unbothered by the league table
- Outcome:
 - Changes in methods for crediting volunteers
 - League table removed (October 2007)
 - Egalitarian methods of credit implemented

http://zoo1.galaxyzoo.org/poster/GalaxyZoo_lowres.jpg



Learning About Volunteers:

2. Serendipitous Discoveries and Green Peas

- Initially, volunteers expected to make classifications only
- Volunteers noticed unusual objects
- One notable *serendipitous discovery* is *Green Peas*
 - Volunteers formed the *Peas Corps*
 - Researchers became involved
 - Peas Corps members rejected offer of article co-authorship

Mon. Not. R. Astron. Soc. 000, 1111–1111 (2009)

doi:10.1111/j.1365-2966.2009.15383.x

Galaxy Zoo Green Peas: discovery of a class of compact extremely star-forming galaxies*

Carolin Caedamone,^{1,2,†} Kevin Schawinski,^{2,3} Marc Sarzi,⁴ Steven P. Bamford,³

Learning About Volunteers:

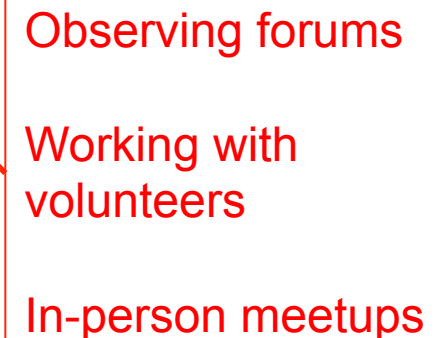
2. Serendipitous Discoveries and Green Peas

- Team members learned about:
 - Volunteers' ability to contribute beyond classification tasks
 - Volunteers' beliefs about legitimate scientific practice

Learning About Volunteers:

2. Serendipitous Discoveries and Green Peas

- Team members learned about:
 - Volunteers' ability to contribute beyond classification tasks
 - Volunteers' beliefs about legitimate scientific practice



Observing forums
Working with
volunteers
In-person meetups

Learning About Volunteers:

2. Serendipitous Discoveries and Green Peas

- Team members learned about:
 - Volunteers' ability to contribute beyond classification tasks
 - Volunteers' beliefs about legitimate scientific practice
- Outcomes:
 - Introduction of *Talk* feature
 - Incorporation into recruitment strategies
 - Launch of new projects for serendipitous discoveries

[www.astro.caltech.edu/
~jwang/PH.png](http://www.astro.caltech.edu/~jwang/PH.png)

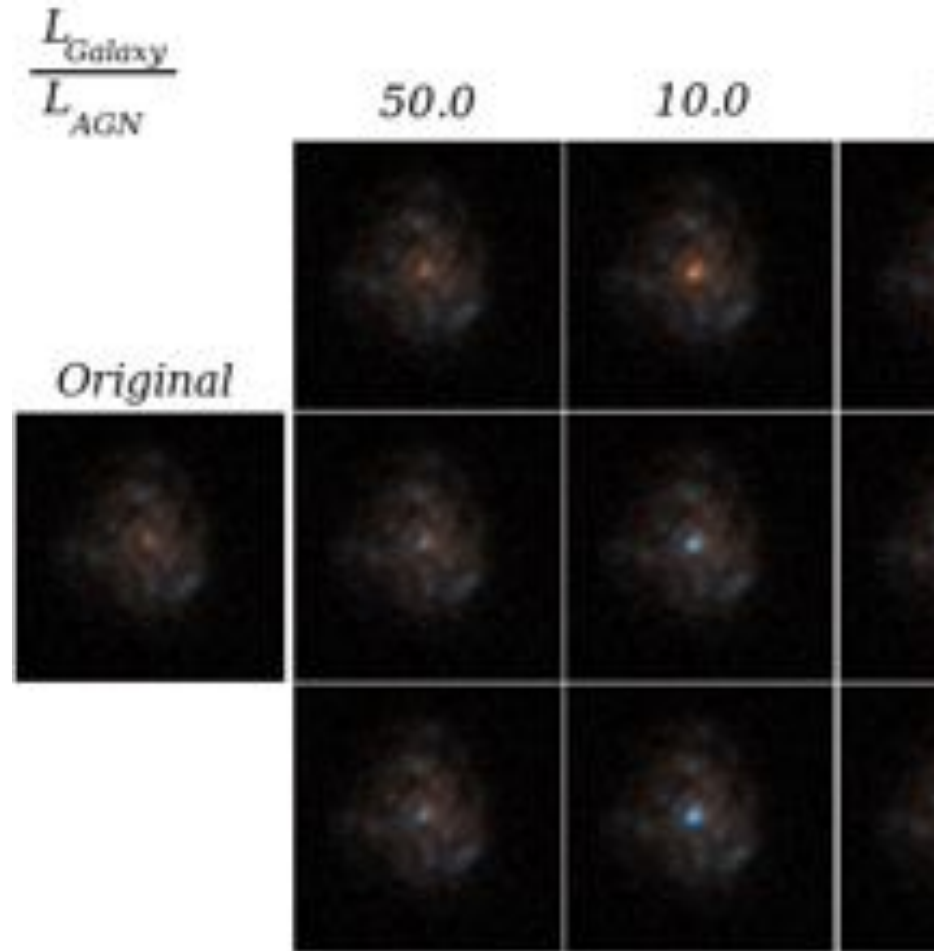
planethunters.org
join the search for exoplanets using the latest Kepler data



Learning About Volunteers:

3. Simulated Images and Bias Testing

- Understanding bias is vital
- Use of simulated images
- Project blogpost described simulated images as “fake”
 - Widespread concern among volunteers
 - One volunteer particularly incited concerns
 - Another volunteer helped to quell concerns



Learning About Volunteers:

3. Simulated Images and Bias Testing

- Team members learned about:
 - Volunteers' beliefs about legitimate scientific practice
 - Dynamics of volunteer community

Learning About Volunteers:

3. Simulated Images and Bias Testing

- Team members learned about:

- Volunteers' beliefs about legitimate scientific practice

Social scientific study of
volunteer motivations

The diagram consists of two red arrows originating from the text box on the right. One arrow points to the bullet point 'Volunteers' beliefs about legitimate scientific practice', and the other points to the bullet point 'Dynamics of volunteer community'.

- Dynamics of volunteer community

Intervening in forums and
comments sections

Learning About Volunteers:

3. Simulated Images and Bias Testing

- Team members learned about:
 - Volunteers' beliefs about legitimate scientific practice
 - Dynamics of volunteer community
- Outcomes:
 - Changes in communication strategies
 - How to reassure volunteers
 - How to use volunteers with strong communication skills

What Did Galaxy Zoo Learn About Volunteers?

- Volunteers' motivations for participation
- Volunteers' abilities to:
 - Engage more deeply with science
 - Influence other volunteers
- Volunteers' view of what is legitimate scientific practice
- Volunteers' interpretations of project features and policies
- Differences between volunteers

How Did Galaxy Zoo Learn About Volunteers?

	More structured	Less structured
Designed to yield insights about volunteers	Social scientific study of volunteer motivations	Sub-forum asking why volunteers participate
Not designed to yield insights about volunteers	Metrics about website usage	Emails from volunteers; In-person meetups; Project forums; Responses to blogposts; Working with volunteers in projects (e.g. Green Peas)

How Did Galaxy Zoo Learn About Volunteers?

	More structured	Less structured
Designed to yield insights about volunteers	Social scientific study of volunteer motivations	Sub-forum asking why volunteers participate
Not designed to yield insights about volunteers	Metrics about website usage	Emails from volunteers; In-person meetups; Project forums; Responses to blogposts; Working with volunteers in projects (e.g. Green Peas)
ADVANTAGES	Greater rigor; Website usage metrics offer rapid feedback	Often provide rapid feedback

How Did Galaxy Zoo Learn About Volunteers?

	More structured	Less structured
Designed to yield insights about volunteers	Social scientific study of volunteer motivations	Sub-forum asking why volunteers participate
Not designed to yield insights about volunteers	Metrics about website usage	Emails from volunteers; In-person meetups; Project forums; Responses to blogposts; Working with volunteers in projects (e.g. Green Peas)
ADVANTAGES	Greater rigor; Website usage metrics offer rapid feedback	Often provide rapid feedback
DISADVANTAGES	Social scientific study is resource-intensive and long-term; Ethical issues of collecting website usage metrics	Prone to bias

Recommendations for Practices

1. Establish methods for aggregating and circulating information about volunteers among team members
2. Make a team member responsible for monitoring forums
3. Regular item in team meetings to discuss volunteers
4. Employ the 'persona method'
5. Seek out ways to increase contact with volunteers:
 1. In-person meetings
 2. Online All-Hands' meetings involving volunteers
6. Give volunteers a seat in decision-making processes
7. Try out new things, but respond quickly if they go wrong

Acknowledgements

- Galaxy Zoo scientists and software engineers
- Galaxy Zoo volunteers
- Doctoral Supervisors: Annamaria Carusi (University of Sheffield) and Marina Jirotko (University of Oxford)
- Doctoral Examiners: David de Roure (University of Oxford) and Robin Williams (University of Edinburgh)
- UCLA Knowledge Infrastructures Team members: Christine Borgman, Milena Golshan, Lo Lee, Irene Pasquetto, Bernie Randles; past members: Sharon Traweek, Jillian Wallis, Ashley Sands, Rebekah Cummings