

Citizen science for demography: new techniques for digital data collection

Evan Roberts, Sociology and Population Studies
University of Minnesota
@evanrobertsnz @MeasuringANZACs

IUSSP Workshop on
Demography in the Digital Age
Wednesday, 25 April 2018

THE SOCIAL SURVEY, ITS HISTORY AND METHODS

By
CARL C. TAYLOR
Associate Professor of Sociology



UNIVERSITY OF MISSOURI
COLUMBIA, MISSOURI

Collaborating with
citizens to research
society has a long
history

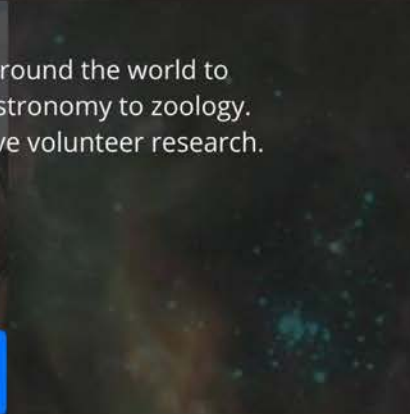
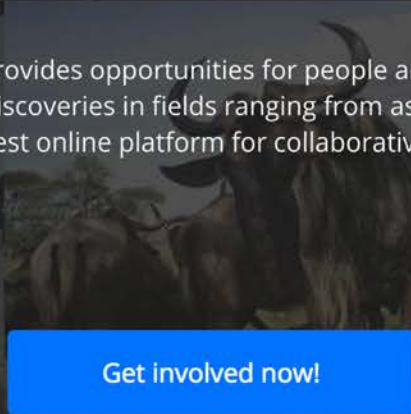
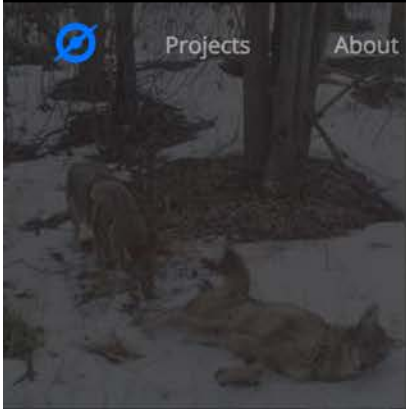
“They have been
largely made by field
workers who do not
have the opportunity
to know and thus
appreciate national
and world situations”

[Projects](#)[About](#)[Talk](#)[Collect](#)[Daily Zooniverse](#)[Blog](#)[BUILD A PROJECT](#)[Sign in](#) [Register](#)

ZOONIVERSE

People-Powered Research

The Zooniverse provides opportunities for people around the world to contribute to real discoveries in fields ranging from astronomy to zoology. Welcome to the largest online platform for collaborative volunteer research.

[Get involved now!](#)

What are these?



And what do they have to do with these?



Who is better at spotting the tiger?



Who is better at spotting the tiger?





WildCam Gorongosa

Identify animals in trail camera images from Gorongosa National Park!

[Get Started](#)



Fossil Finder

Join us in the search and discovery of fossils at Lake Turkana, Kenya.

[Get Started](#)



Galaxy Zoo: Bar Lengths

Measure the engines of evolution in disk galaxies.

[Get Started](#)



Whales as Individuals

Help us identify individual Humpback Whales by clueing our computer algorithms in to patterns on their tails

[Get Started](#)



Season Spotter Image Marking

Help keep an eye on changing seasons by marking images!

[Get Started](#)



Season Spotter Questions

Help keep an eye on changing seasons by answering questions!

[Get Started](#)



Chimp & See

Discover the secret life of chimpanzees. We need your help to study, explore, and learn from thousands of videos.

[Get Started](#)



AnnoTate

Help transcribe documents from the Tate collection, and reveal the secret lives of artists.

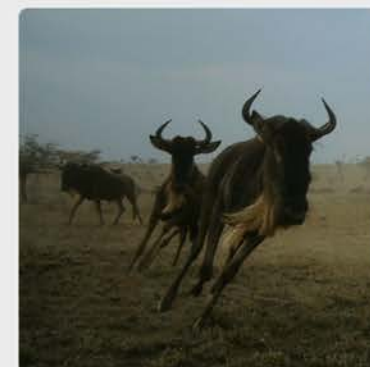
[Get Started](#)



Science Gossip

Uncover the history of citizen science. Help us to classify their drawings and map the origins of citizen science.

[Get Started](#)



Wildebeest Watch

Explore collective intelligence in wildebeest!

[Get Started](#)

Snapshot Serengeti is a Zooniverse project.

The Zooniverse is a collection of web-based Citizen Science projects that use the efforts and abilities of volunteers to help researchers deal with the flood of data that confronts them.

Our Projects

We currently have 12 projects on subjects ranging from [astronomy](#), to [climatology](#), to [biology](#), to [humanities](#).

[Forgot Password?](#)

English

SNAPSHOT SERENGETI

[Home](#)[About](#)[Classify](#)[Profile](#)[Discuss](#)[Blog](#)[Authors](#)

Welcome to Snapshot Serengeti

Hundreds of camera traps in Serengeti National Park, Tanzania, are providing a powerful new window into the dynamics of Africa's most elusive wildlife species. We need your help to classify all the different animals caught in millions of camera trap images.

Season 1

Season 2

Season 3

Season 4

Season 5

Season 6

Season 7

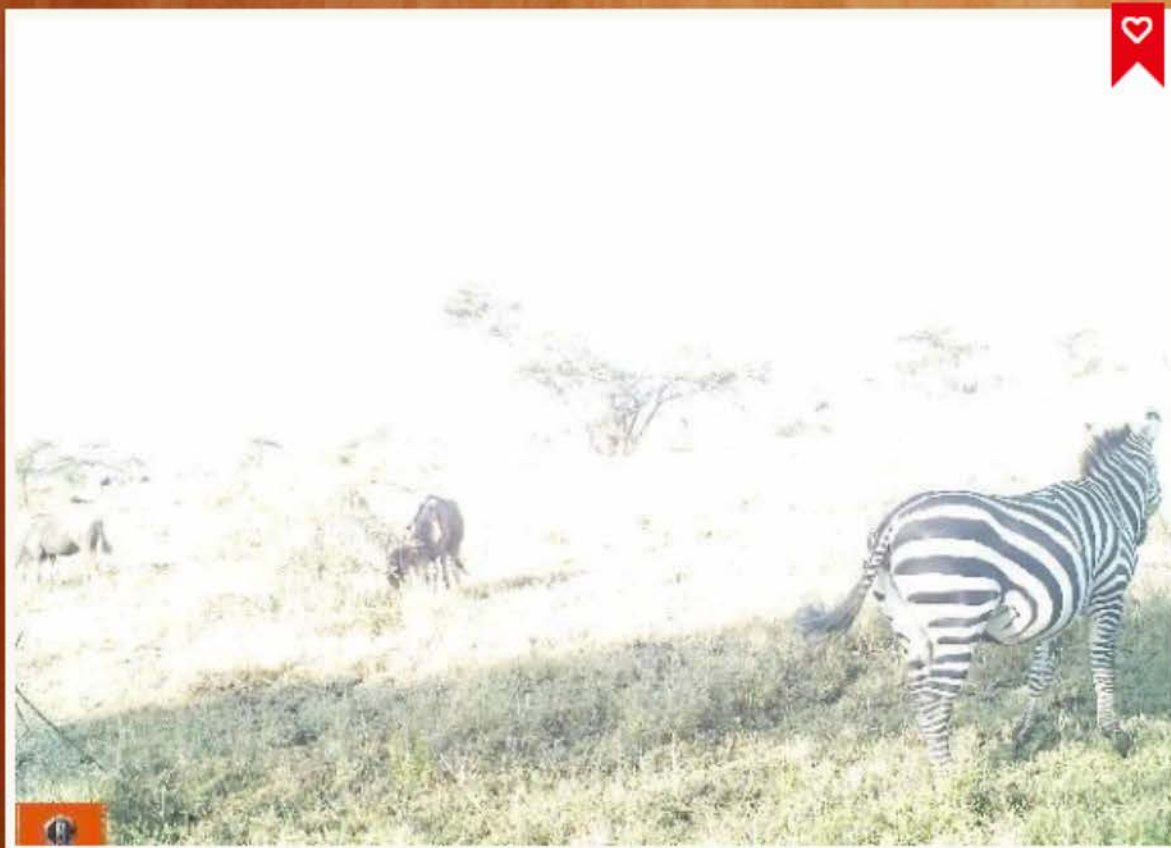
Season 8

Lost Season

Season 9

Season 9.5

[Start classifying](#)


☐ Fire

☐ No animals present

Finish

Looks like ▾



Pattern

Color

Horns

Tail

Build

Aardvark

Genet

Porcupine

Aardwolf

Giraffe

Reedbuck

Baboon

Guinea fowl

Reptiles

Bat

Hare

Rhinoceros

Bat-eared fox

Hartebeest

Rodents

Bird (other)

Hippopotamus

Secretary bird

Buffalo

Honey-badger

Serval

Bushbuck

Hyena (spotted)

Steenbok

Cattle

Hyena (striped)

Topi

Caracal

Impala

Vervet monkey

Cheetah

Insect/Spider

Vulture

Civet

Jackal

Warthog

Dik dik

Kori bustard

Waterbuck

Duiker

Leopard

Wildcat

Eland

Lion (female or cub)

Wildebeest

Elephant

Lion (male)

Zebra

Gazelle (Grant's)

Mongoose

Zorilla

Gazelle (Thomson's)

Ostrich

Human

Citizen science is good for research situations where:

1. The volume of digitized raw data, particularly text, images, videos and sound has increased significantly;
2. Computers currently do a poor job of classifying this data;
3. Citizens with a small amount of training can produce reasonably accurate judgments;
4. Combining the insights of multiple citizens yields measurements as accurate as produced by a subject expert

Types of citizen science research

1. Image, video and sound classification
 - ❖ Collaborative ethnographic interpretation
 - ❖ Post-disaster research
2. Transcription and text recognition
3. Citizen data gathering. Environmental monitoring [broadly construed] – used in ecology. Smartphones and other devices enhance possibilities of citizen data capture. VGIS and other volunteered information.

Demographic application 1.
Transcription of complex material

Measuring the ANZACs
<http://www.measuringtheanzacs.org/>
@MeasuringANZACs

US Colored Troops
<http://usct.cc>
@transcribe_usct

Measuring the ANZACs

Do your part to help transcribe first-hand accounts of New Zealanders from the Australian and New Zealand Army Corps circa World War I.

[START MARKING](#)

or

[START TRANSCRIBING](#)

SCRIBE

This project is built using Scribe: document transcription, crowdsourced.

What kind of page is this? ?

HISTORY SHEET

STATEMENT OF SERVICES

ATTESTATION: GENERAL FORM

ATTESTATION: MEDICAL FORM

DEATH NOTIFICATION

SOUTH AFRICAN WAR ATTESTATION

OTHER

DONE

BAD PAGE

1. Classify pages to
open up
page-dependent
menus and instructions

2. Citizen scientists identify fields and mark their location for transcription

Mark is used to identify specific pieces of data; for example, dates, nationalities, heights and weights, etc. You will see marks submitted by other volunteers, although you can hide these for visibility. If an item is already marked, you can skip over it! You can also delete others' marks, which counts as a vote for that mark being incorrect.

Select a label and draw rectangle around any corresponding on the document.

Unit (7) ?

Rank (7)

Surname (6)

Christian name (8)

No. (7)

Occupation (6)

Last Employer (6)

Religion (6)

Last New Zealand Address (8)

☐ Next-of-Kin (4)

NEXT

The image shows a historical military service form with several fields marked by blue rectangles. The fields are:

- Unit (7)
- Rank (7)
- Surname (6)
- Christian name (8)
- No. (7)
- Occupation (6)
- Last Employer (6)
- Religion (6)
- Last New Zealand Address (8)
- Next-of-Kin (4)

The form also includes sections for:

- Service
- Periods of Service
- Theatres of Operation

Handwritten entries include:

- No. 23/334
- Rank: 2/Lt
- Name: A. L. C. Cornick
- Address: Waiuku
- In New Zealand: 192 days
- Overseas: 3 years 104 days
- Total service: 3 years 296 days
- Date commenced duty: 28/1/15
- Theatres of Operation: Australasian, Egyptian 1915-16, Egyptian E.F. 1916, Balkan, Western 1916-17

Y-SHEET. Form No. 3A.

Authority to Headmaster at Table.

Christian Name: Aloysius

No: 5/227

James Aloysius John

7/227

Religion: Roman

Last New Zealand address: Raikoura

6 MAR 1952

Mackie (father)

ough.

Mackie

Raikoura

Theatres of Operation.

What kind of page is this?

HISTORY SHEET

STATEMENT OF SERVICES

ATTESTATION: GENERAL FORM

ATTESTATION: MEDICAL FORM

DEATH NOTIFICATION

SOUTH AFRICAN WAR ATTESTATION

OTHER

DONE

BAD PAGE

[View A Tutorial](#)

☒ Transcribe this page now!

☐ Discuss this personnel record.



Show All

3.
Transcribe
marked
fields via two
routes ...

Measuring the ANZACs

Do your part to help transcribe first-hand accounts of New Zealanders from the Australian and New Zealand Army Corps circa World War I.

[START MARKING](#)

or

[START TRANSCRIBING](#)

SCRIBE

This project is built using Scribe: document transcription, crowdsourced.

Essentially random entries not yet transcribed

Rank	Surname	Christian Name	No.
TRANSCRIBE	ABOUT	SEARCH	FIELD GUIDE
BLOG	DISCUSS	HELLO EVAN ROB	
PVTE	Macaw	James Courtney	38196
Religion:		Last New Zealand address:	
Ireland		Hastings	
C.E.			
21-12-80			
Mrs K. Macaw (Mother)		T.G. Da	
Innistrush.			
Ireland.			
No. 38196		Rank: Pte	
Name: James Courtney Macaw		Days.	

Transcribe last New Zealand address. Press [Shift] + [Enter] to start a new line.

[Need some help?](#)[Bad mark](#)[Illegible?](#)[Next Entry](#)

Accuracy of RAs versus citizens?!

	Attestations (NZ RAs)		Casualty Rolls (Canadian RAs)	
	Given	Last	Given	Last
Mean similarity score to truth (1 = absolute accuracy)	0.98	0.99	0.99	0.99
Proportion absolutely accurate	0.91	0.95	0.93	0.98
Proportion with similarity score > 0.9	0.95	0.98	0.95	0.99
Proportion with similarity score > 0.8	0.97	0.99	0.99	0.99
Proportion with > 3 words in string (more likely problematic transcriptions)	0.003	0.0002	0.030	0
Citizen transcription				
	Given name	Surname	Serial number	
Mean similarity score to truth (1 = absolute accuracy)	0.98	0.98	0.98	
Proportion absolutely accurate	0.87	0.84	0.89	
Proportion with similarity score > 0.9	0.95	0.95	0.95	
Proportion with similarity score > 0.8	0.97	0.98	0.98	
Proportion with > 3 words in string (more likely problematic transcriptions)	0.003	0.003	0.003	



African American Civil War Soldiers

[ABOUT](#)[CLASSIFY](#)[TA](#)

Me | **8** | **U.S.C.T.**

Jordan M. Craig

Co. *16*, 8 Reg't U. S. Col'd Inf.

Appears on

Company Descriptive Book

of the organization named above.

DESCRIPTION.

Age *18* years height *5* feet *5 1/2* inches.

Complexion *Blk*

Eyes; *Blk* hair *Blk*

Where born *Indiana*

Occupation *Hostler*

ENLISTMENT.

When *Sept 28, 1864*

Where *Indianapolis*

By whom *Capt Braden*; term *1* y'r.

Remarks: *Sub. Mustered out at Ringgold Barracks Texas. Expiration term of service*

Name:

Co.:

Company (see field guide for letters)

Select...

Age:

Height:

Feet

Select...

Inches

N/A

Inch fractions (optional)

N/A

[Need some help with this task?](#)[Back](#)[Next](#)



African American Civil War Soldiers

[ABOUT](#)[CLASSIFY](#)[TA](#)

Me | **8** | **U.S.C.T.**

Jordan M. Craig

Co. *16*, 8 Reg't U. S. Col'd Inf.

Appears on

Company Descriptive Book

of the organization named above.

DESCRIPTION.

Age *18* years height *5* feet *5 1/2* inches.

Complexion *Blk*

Eyes; *Blk* hair *Blk*

Where born *Indiana*

Occupation *Hostler*

ENLISTMENT.

When *Sept 28, 1864*

Where *Indianapolis*

By whom *Capt Braden*; term *1* y'r.

Remarks: *Sub. Mustered out at Ringgold Barracks Texas. Expiration term of service*

Name:

Co.:

Company (see field guide for letters)

Select...

Age:

Height:

Feet

Select...

Inches

N/A

Inch fractions (optional)

N/A

[Need some help with this task?](#)[Back](#)[Next](#)

Demographic application 2: Image classification

Earthworm Invasion

Earthworms are causing major changes in hardwood forests



Earthworm affected



No worms

z.umn.edu/earthworms

The screenshot shows the website's layout. At the top is a dark navigation bar with links: PROJECTS, ABOUT, GET INVOLVED, TALK, BUILD A PROJECT, NEWS, NOTIFICATIONS, MESSAGES, and RUNCK014. Below this is a secondary navigation bar with a circular profile picture and links: EARTHWORM INVASIONS! CLASSIFYING HISTORICAL LAND USES, ABOUT, CLASSIFY, TALK, COLLECT, RECENTS, and LAB. The main content area features a large background image of a field with the text "Help us understand how land use impacts earthworm invasion" and a "Learn more" button. Below this, a "Get started" section with a downward arrow provides instructions and four buttons: "Classify forests and trees", "Classify roads", "Classify agriculture", and "Identify buildings and farm yards".

PROJECTS ABOUT GET INVOLVED TALK BUILD A PROJECT NEWS NOTIFICATIONS MESSAGES RUNCK014

EARTHWORM INVASIONS! CLASSIFYING HISTORICAL LAND USES ABOUT CLASSIFY TALK COLLECT RECENTS LAB

Help us understand how land use impacts earthworm invasion

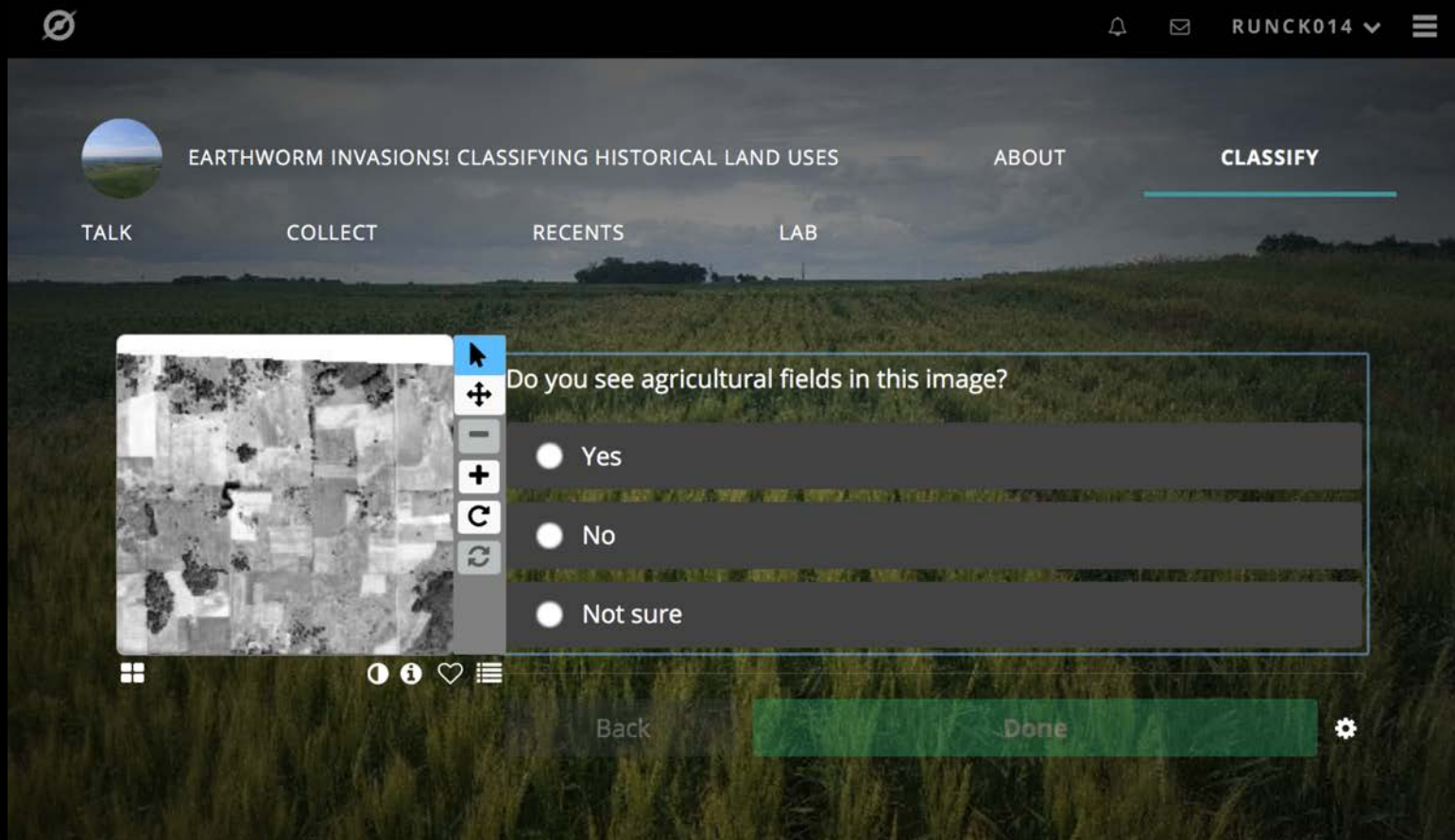
Learn more

Get started ↓

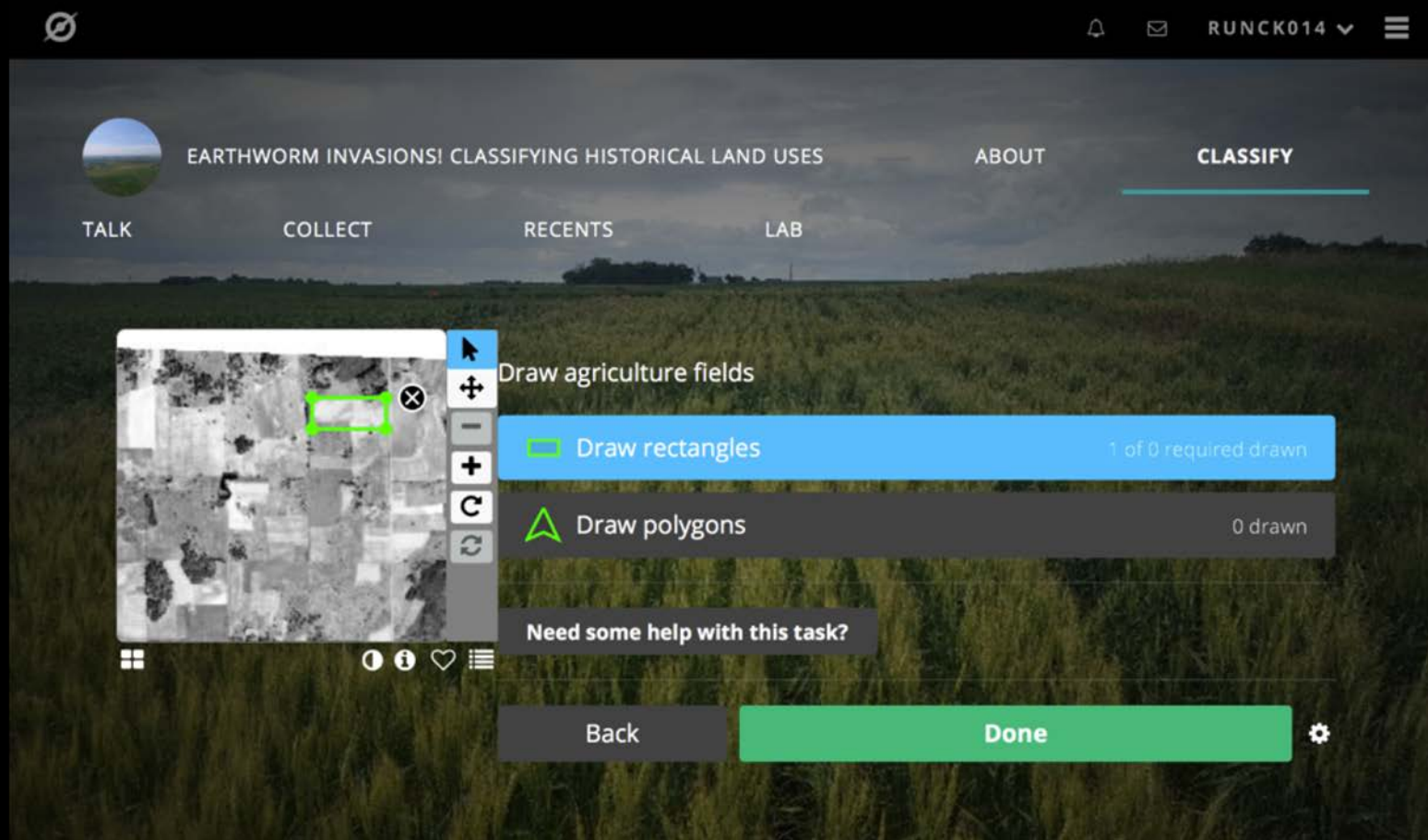
There are multiple options for how to help classify land use. Each task involves drawing. 1) draw roads using lines 2) draw farm fields using rectangles 3) draw in forest using rectangles 4)

Classify forests and trees Classify roads Classify agriculture Identify buildings and farm yards

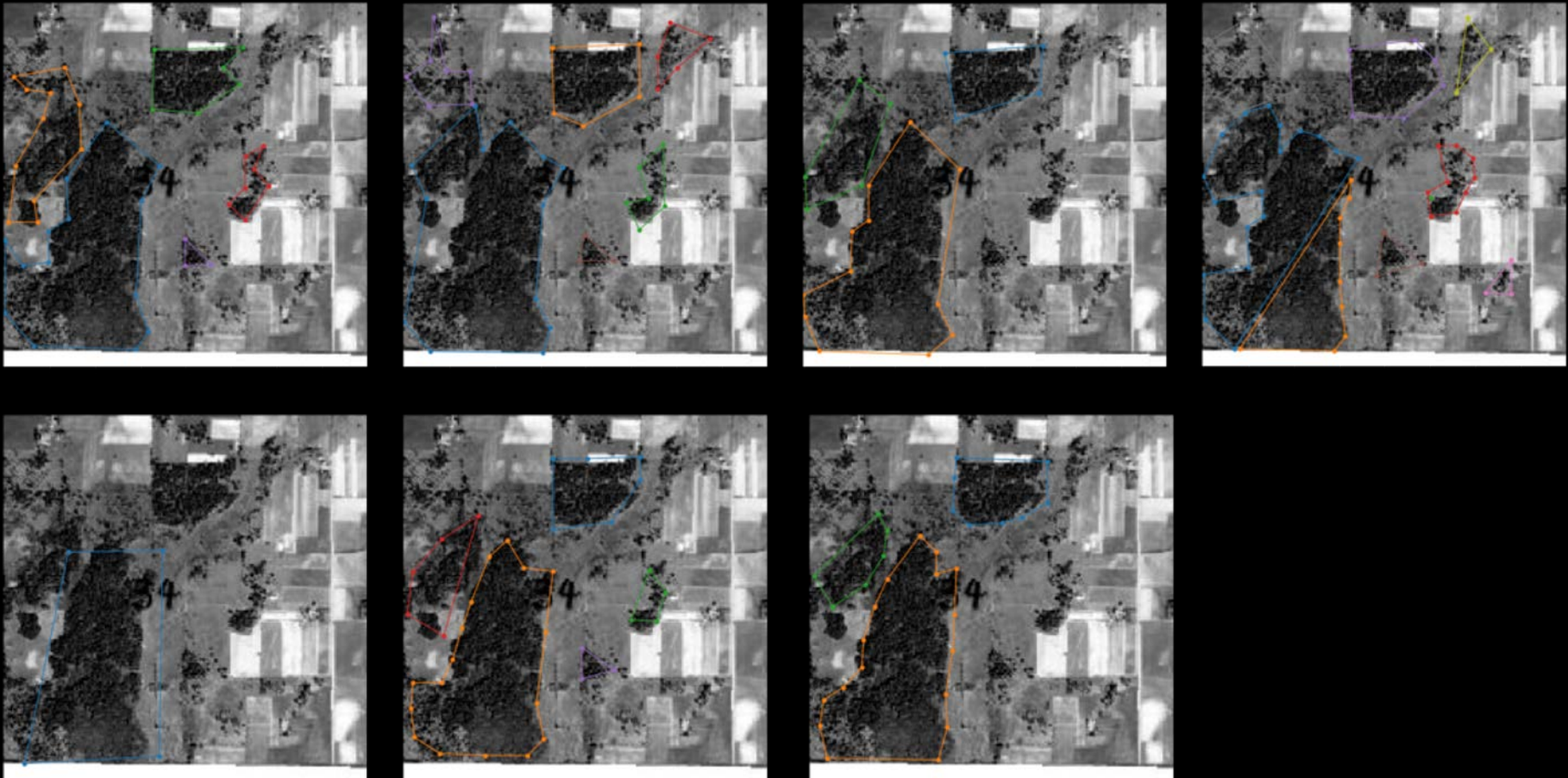
Step 1: Do you see a field?



Step 2: Place polygons around the fields

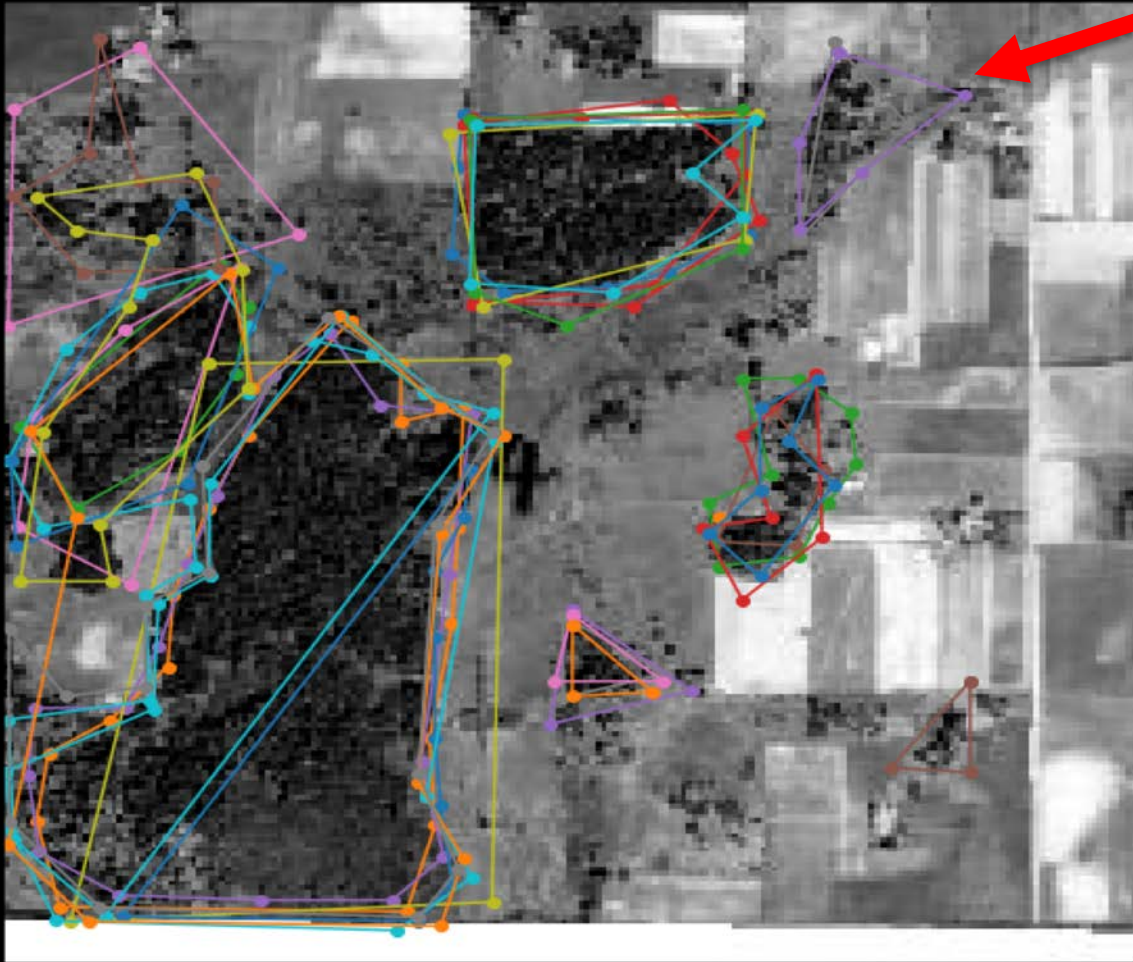


Humans 1-7: Borgholm 1939 image 3



n=7: 24.00 % Forest on average
(17.95% to 31.09%)

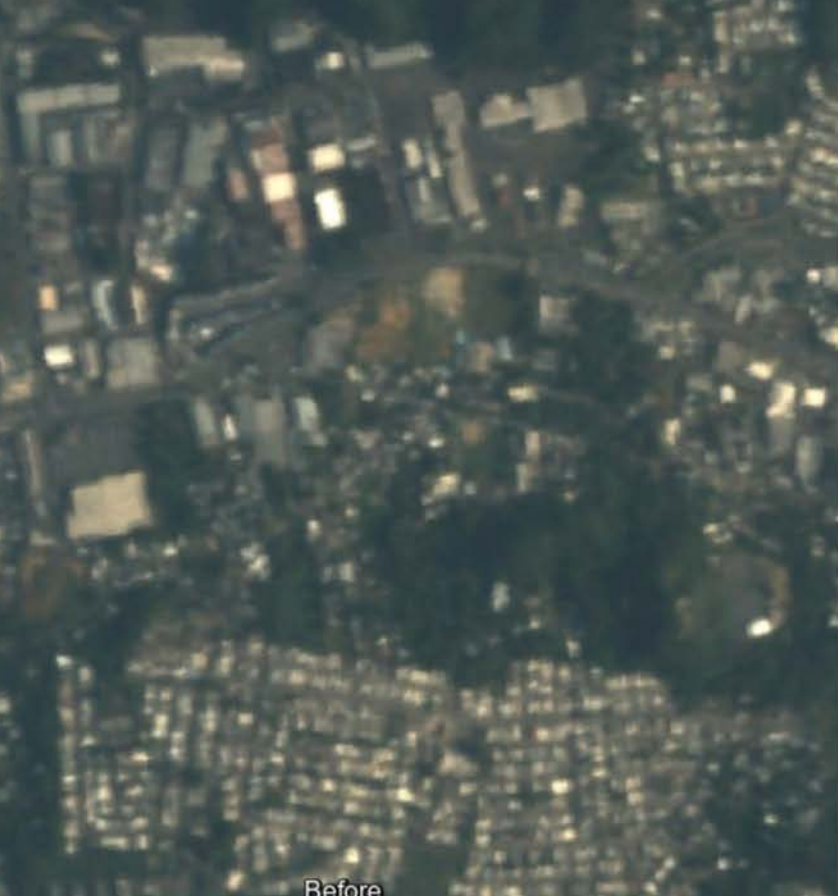
Humans 1-7: Borgholm 1939 image 3



n=7
24.00 % Forest
on average
(17.95% to 31.09%)

Disaster response: Image classification

FINISHED!



Before

Mark the features you see.
If you don't see any, just click "Done" to move on.

If you're not sure what to do, click the "Need some help with this task?" button below.

Road Blockage0 drawn

Flood0 drawn

Temporary Settlement0 drawn

Structural damage0 drawn

Need some help with this task?

☐ Hide previous marks

Unclassifiable Image

Ocean Only (no land)

Done & Talk

Done

Show the project tutorial