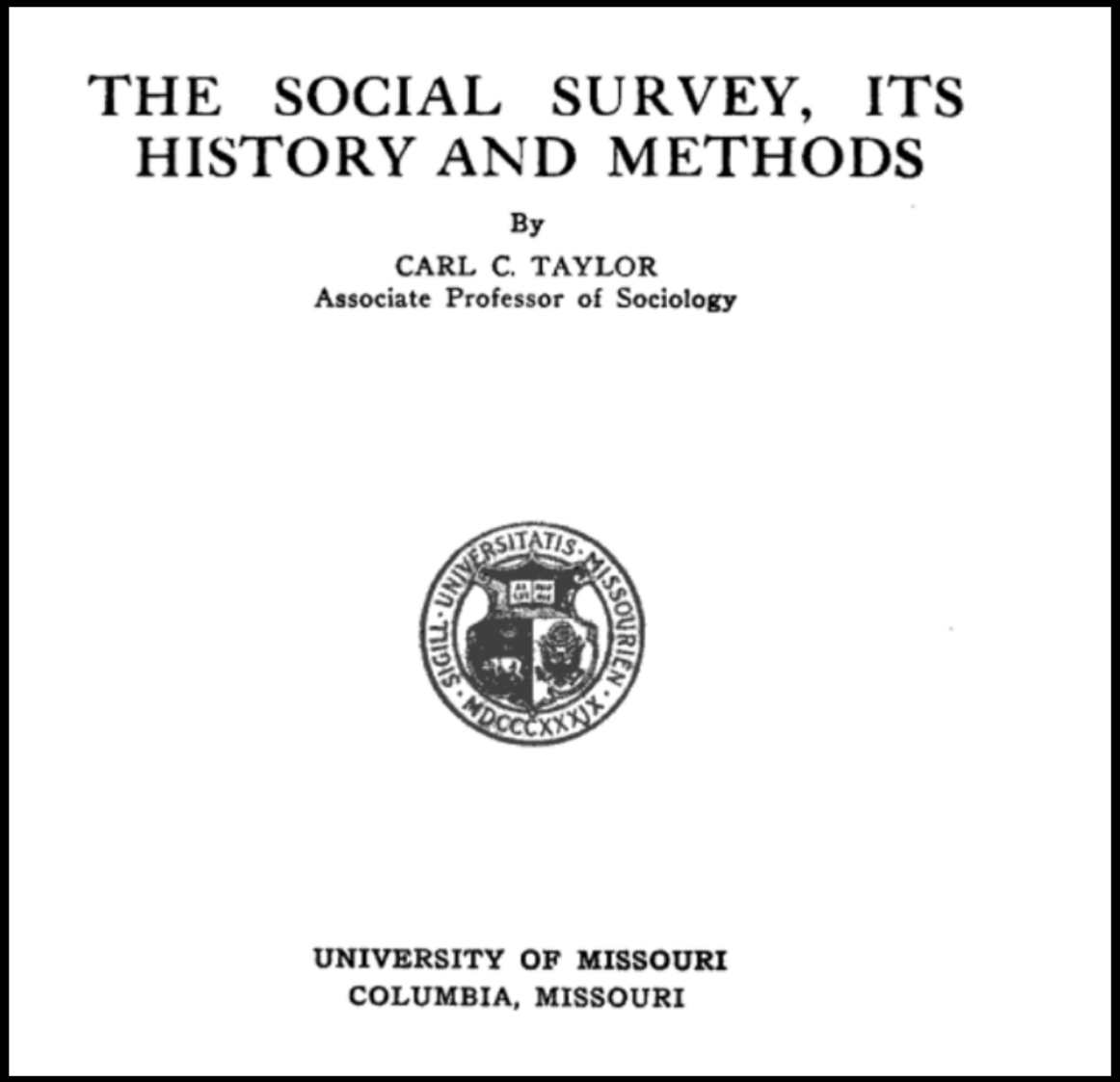
Citizen science for demography: new techniques for digital data collection

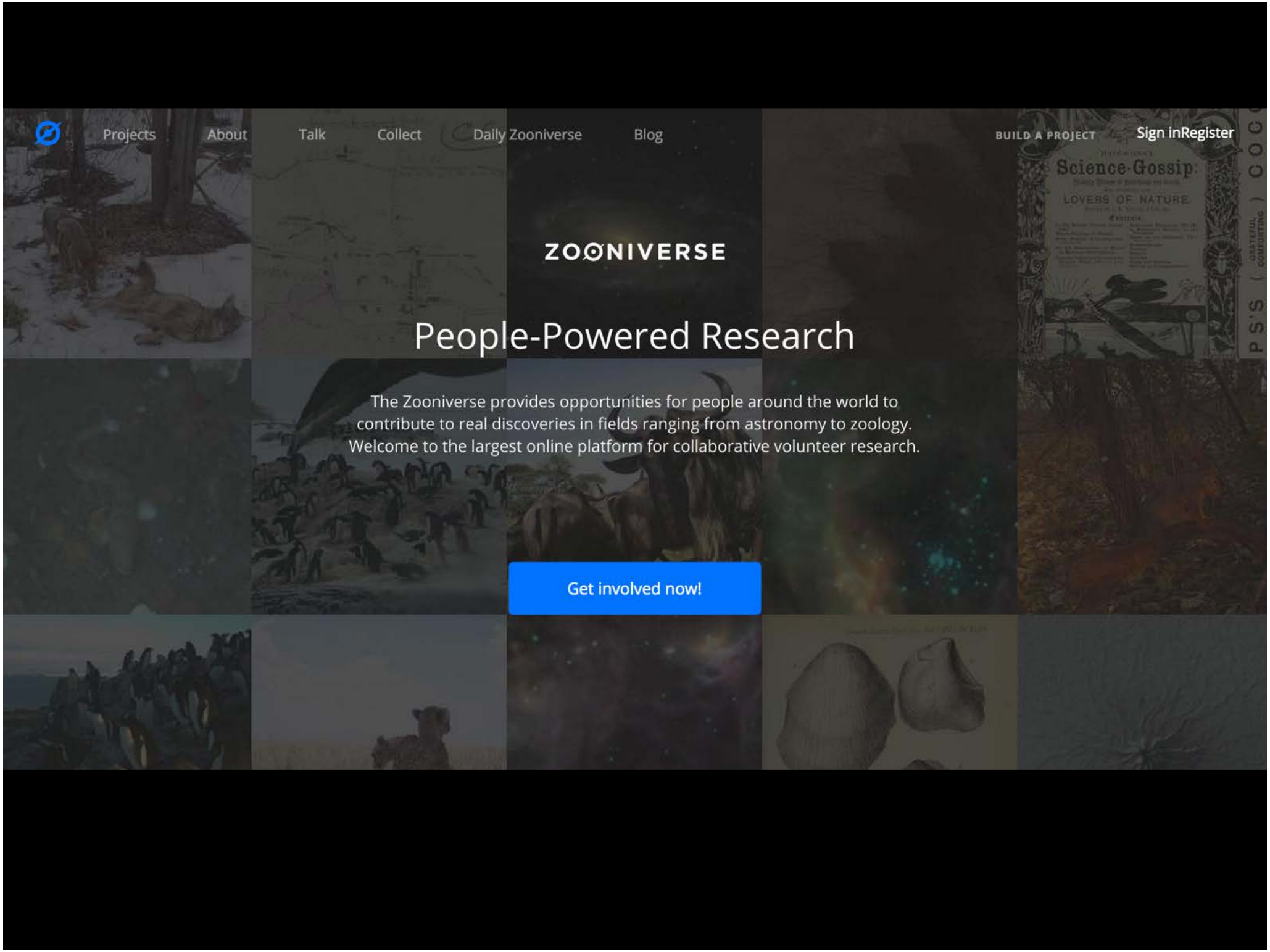
Evan Roberts, Sociology and Population Studies
University of Minnesota
@evanrobertsnz @MeasuringANZACs

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



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 inRegister
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

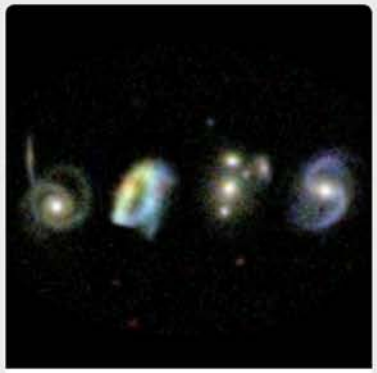
Identify animals in trail camera images from Gorongosa National Park!

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 reseachers 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.

username
Forgot Password?

English

SNAPSHOT SERENGETI

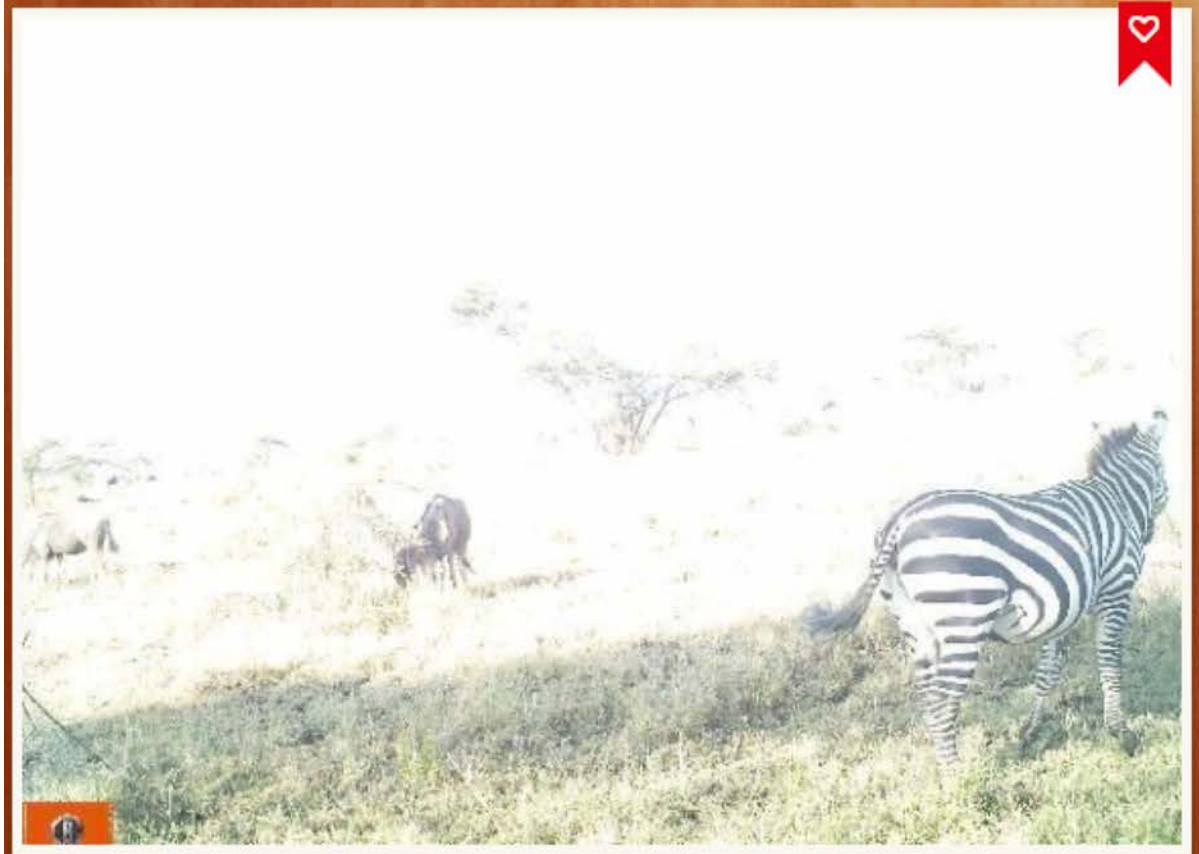
Home | About | Classify | Profile | Discuss | Blog | Authors

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

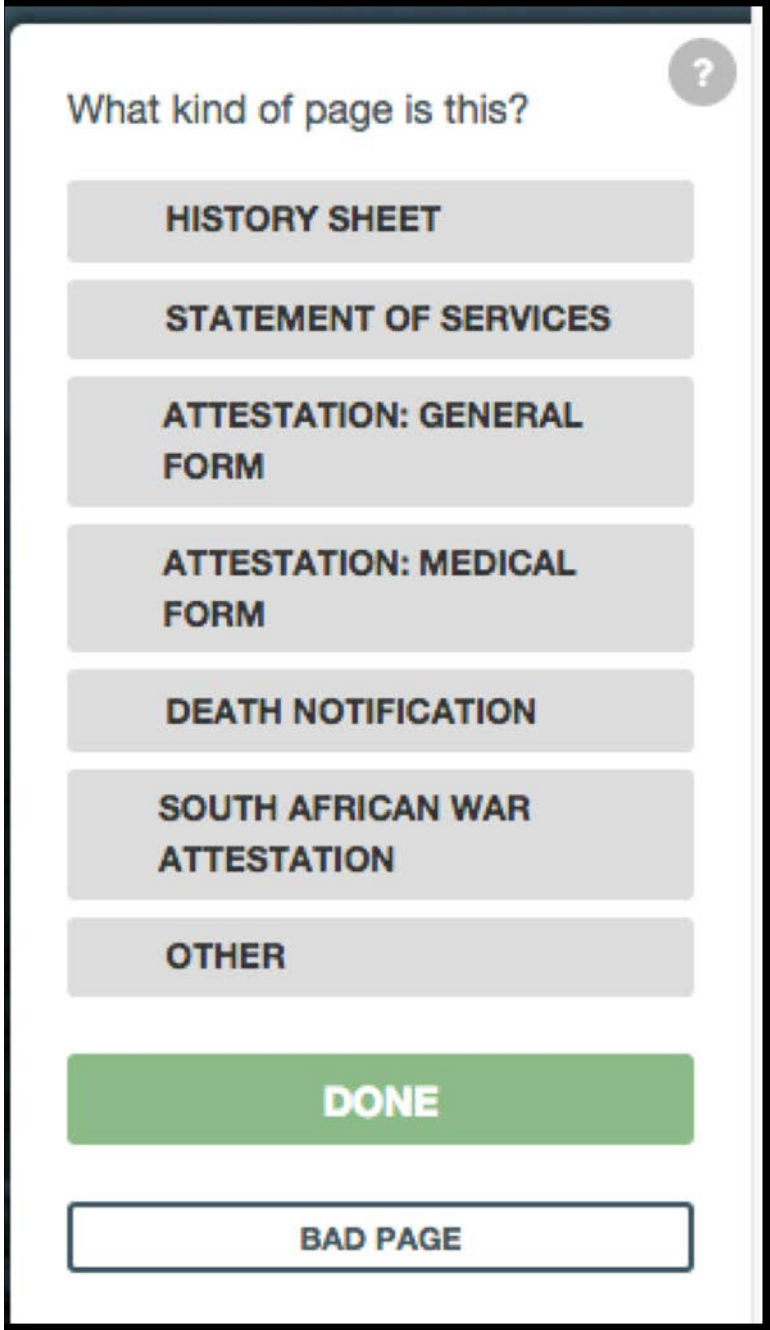
1. Image, video and sound classification
   - ❖ Collaborative ethnographic interpretation
   - ❖ Post-disaster research
2. Transcription and text recognition
3. Citizen data gathering. Environmental monitoring [broadly construed] – used inecology. Smartphones and other devices enhance possibilities of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

Demographic application 1. Transcription of complex material

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

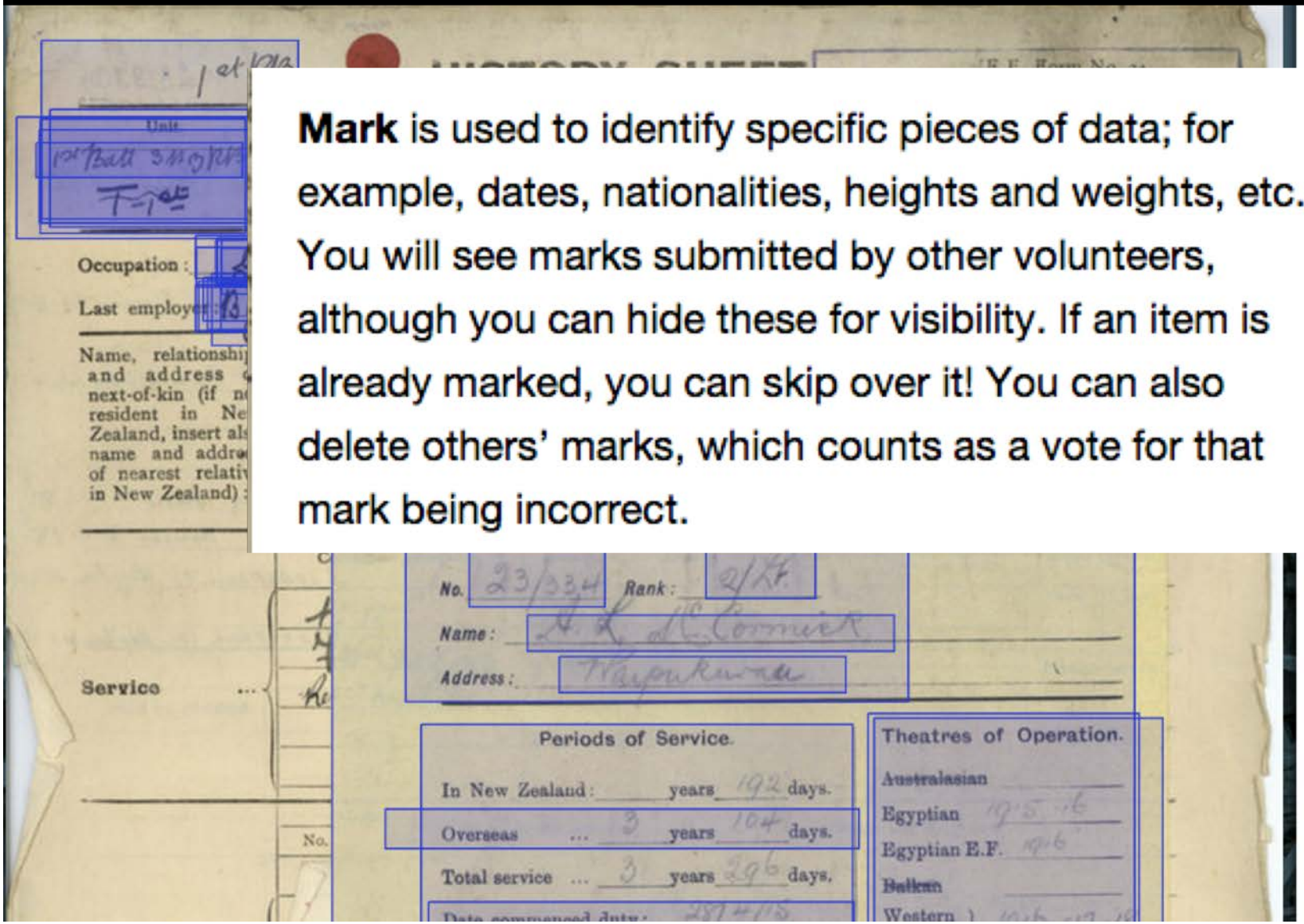
US Colored Troops
http://usct.cc
@transcribe_usct

ANZAC
MARK TRANSCRIBE ABOUT FIELD GUIDE BLOG DISCUSS
HELLO EVAN ROBERTS, MEASURING THE ANZACS LOGOUT
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.

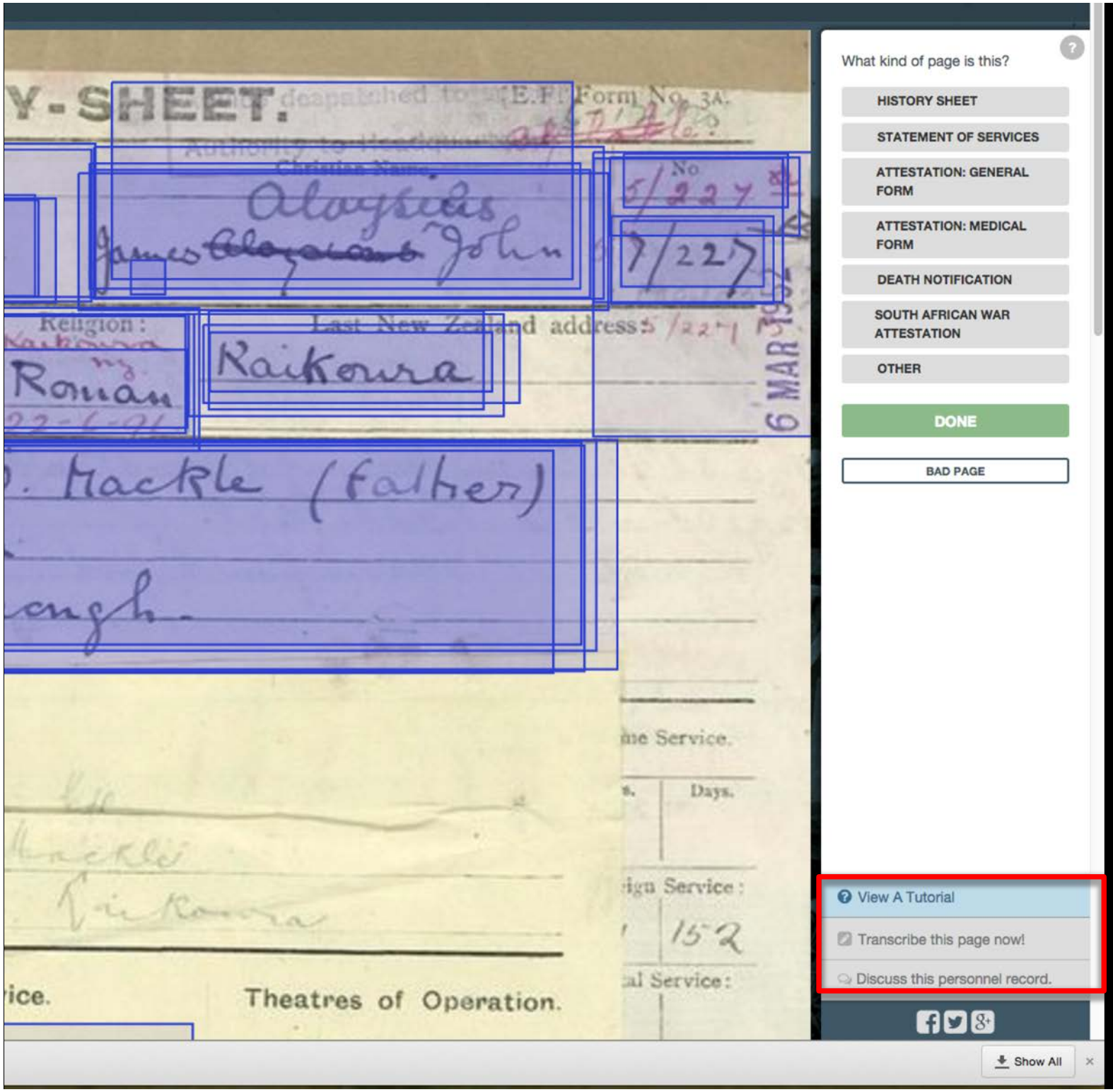


1. Classify pages to open up page-dependent menus and instructions

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



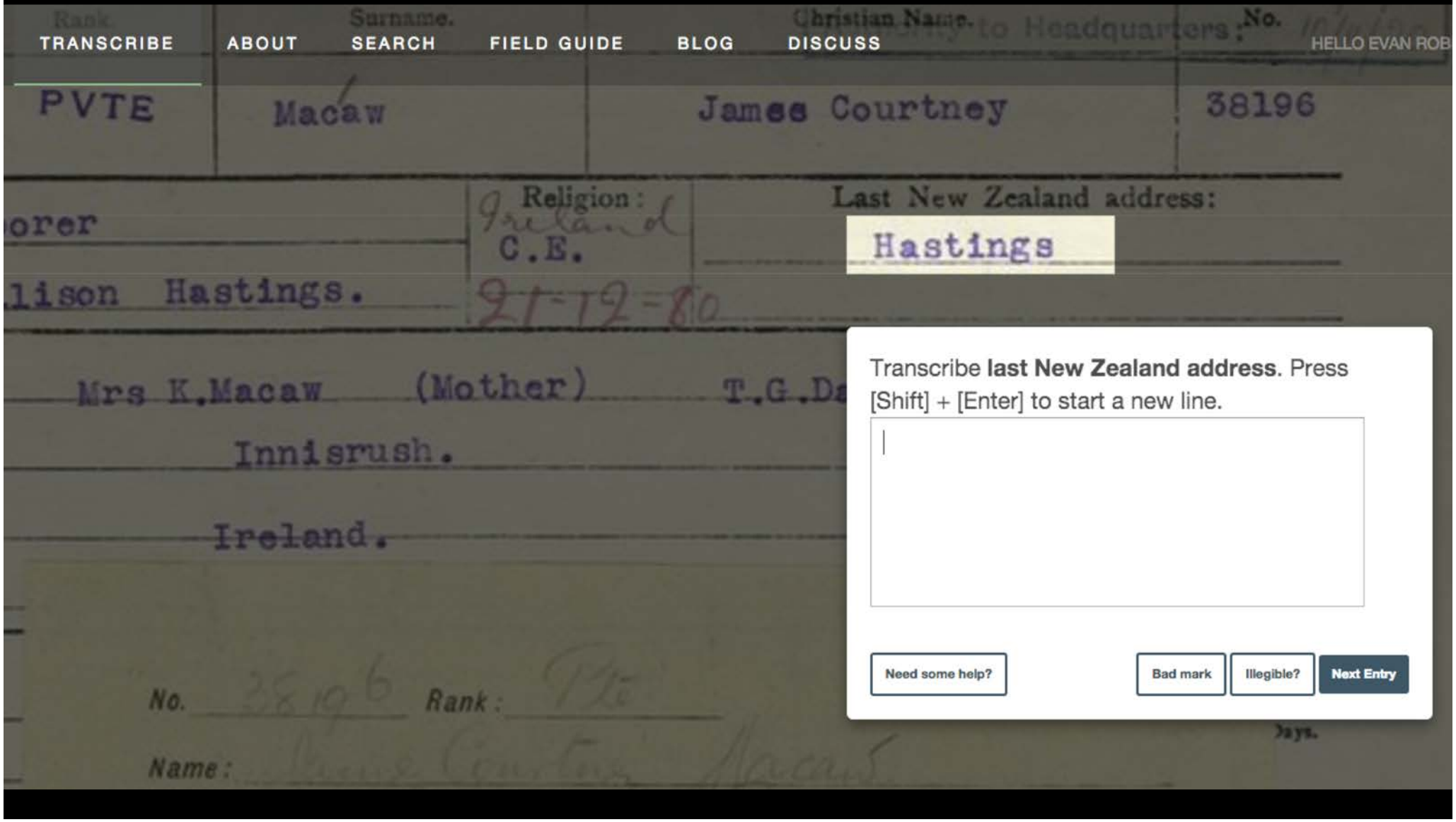
Mark is used to identify specific pieces of data; for example, dates, nationalities, heights and weights, etc. You will see marks submitted by other volunteers, although you can hide these for visibility. If an item is already marked, you can skip over it! You can also delete others' marks, which counts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



3.
Transcribe marked fields via two routes …

ANZAC
MARK TRANSCRIBE ABOUT FIELD GUIDE BLOG DISCUSS
HELLO EVAN ROBERTS, MEASURING THE ANZACS LOGOUT
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.

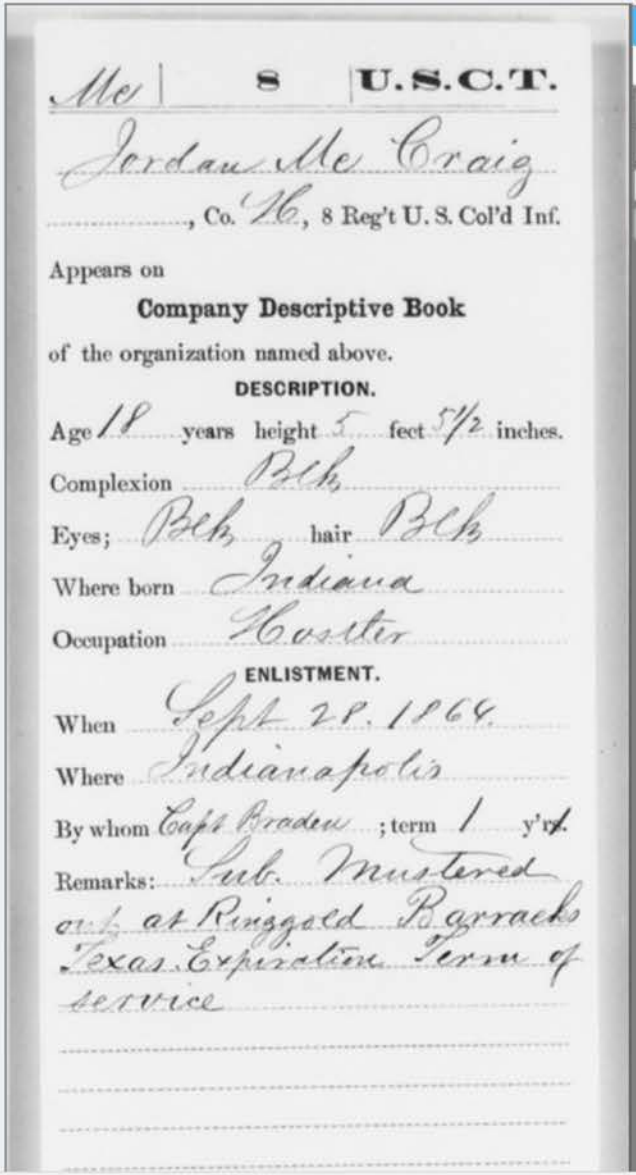
Essentially random entries not yet transcribed



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.	0.98	0.95	0.99
Proportion with similarity score > 0.8		0.99	0.99	0.99
Proportion with > 3 words in string (more likely problematic transcriptions)		0.0002	0.030	0

	Citizen transcription		
	Given ne	Surname	Serial number
Mean similarity score to truth (1 = absolute accuracy)	98	0.98	0.98
Proportion absolutely accurate	.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



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

Jordan Mc 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

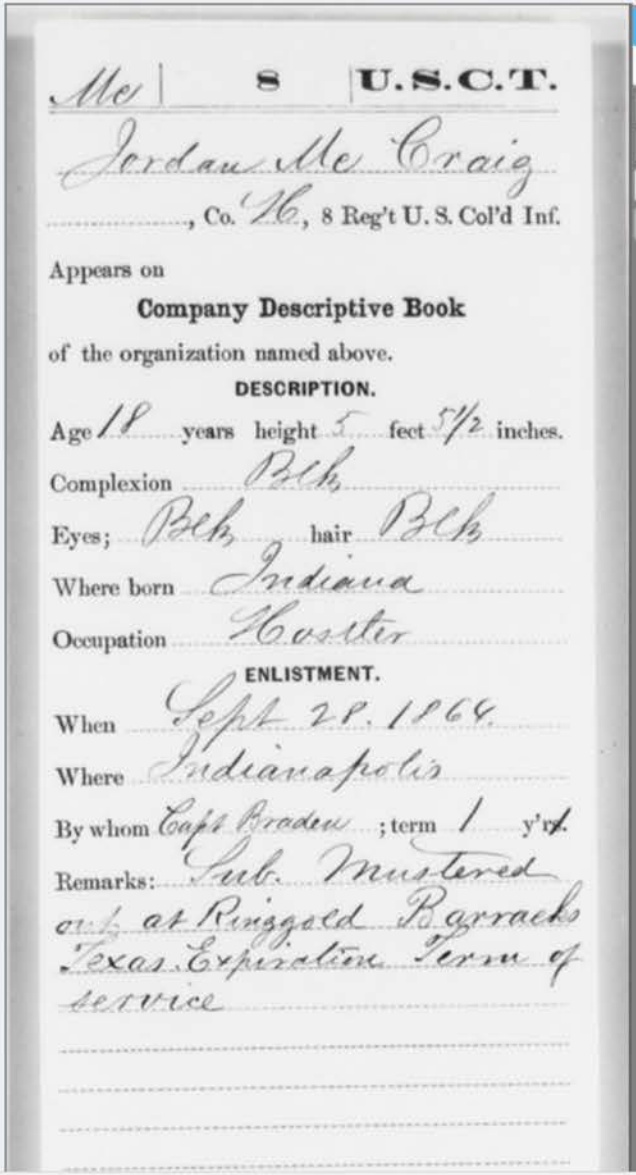
N/A

Inch fractions (optional)

N/A

Need some help with this task?

Back Next



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

Jordan Mc 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

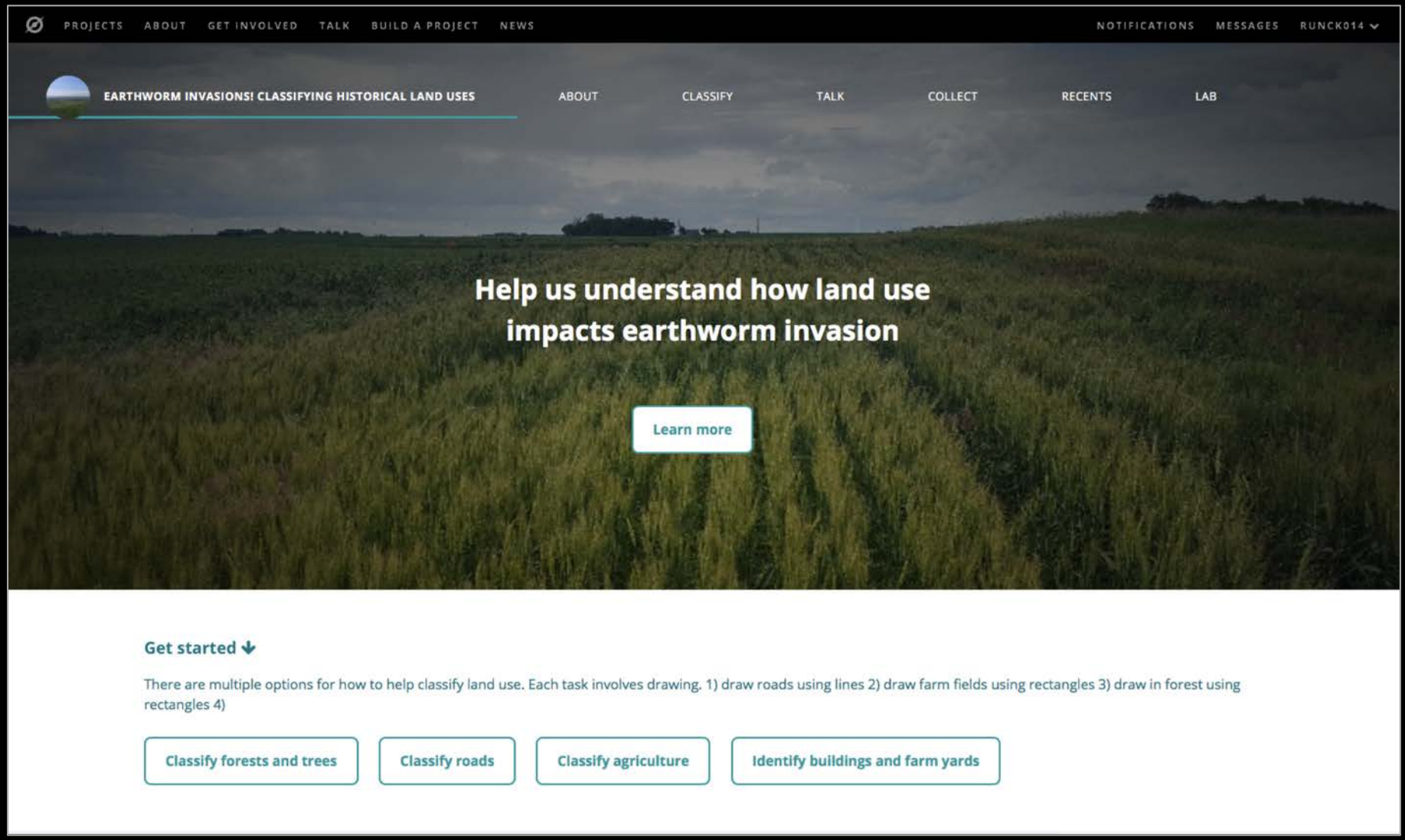
Earthworms are causing major changes in hardwood forests

Earthworm affected

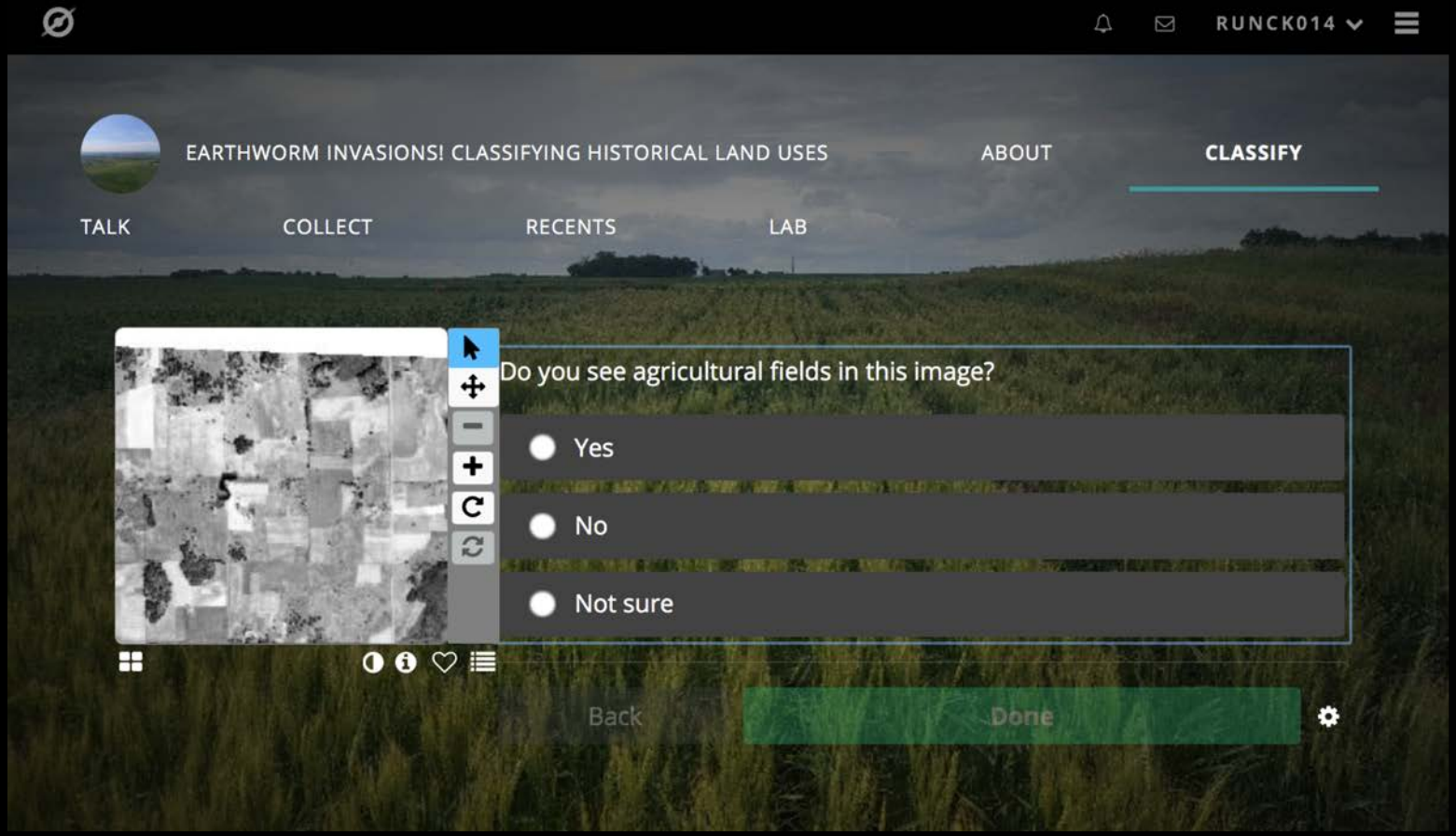
No worms

Photo credit: Smithsonian Insider
Craven et al. 2016. Global Change Biology

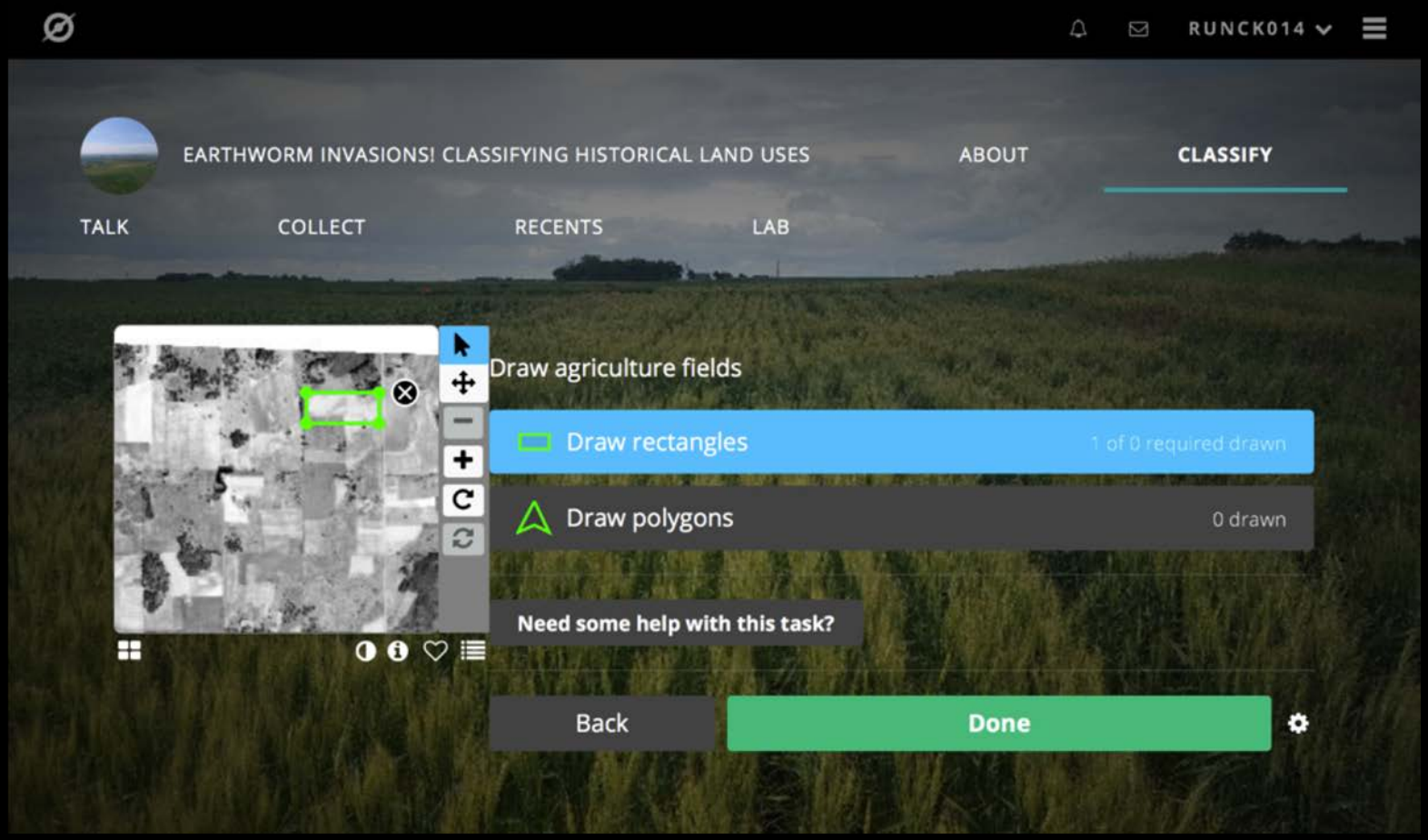
z.umn.edu/earthworms



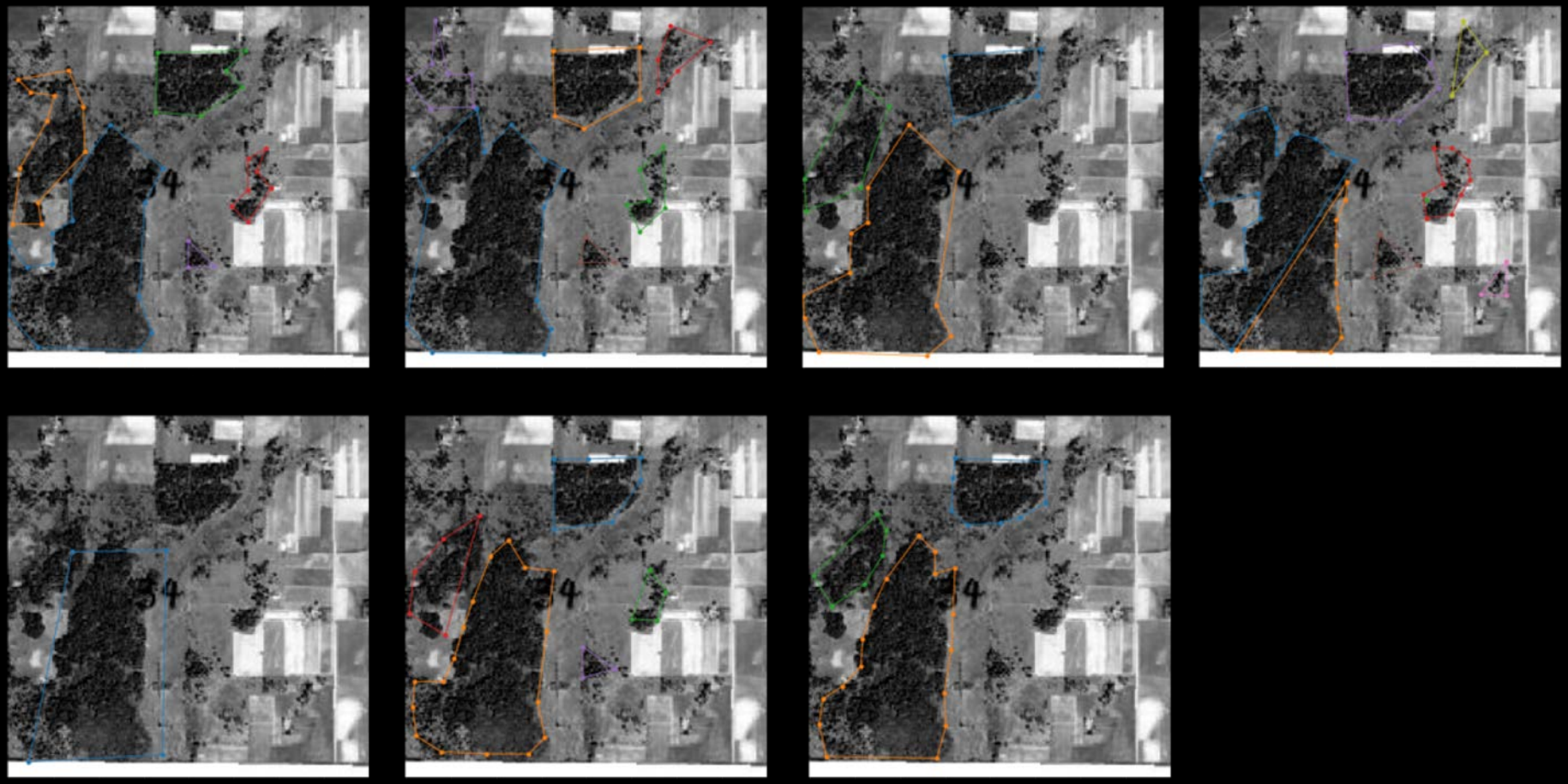
Step 1: Do you see a field?



Step 2: Place polygons around the fields

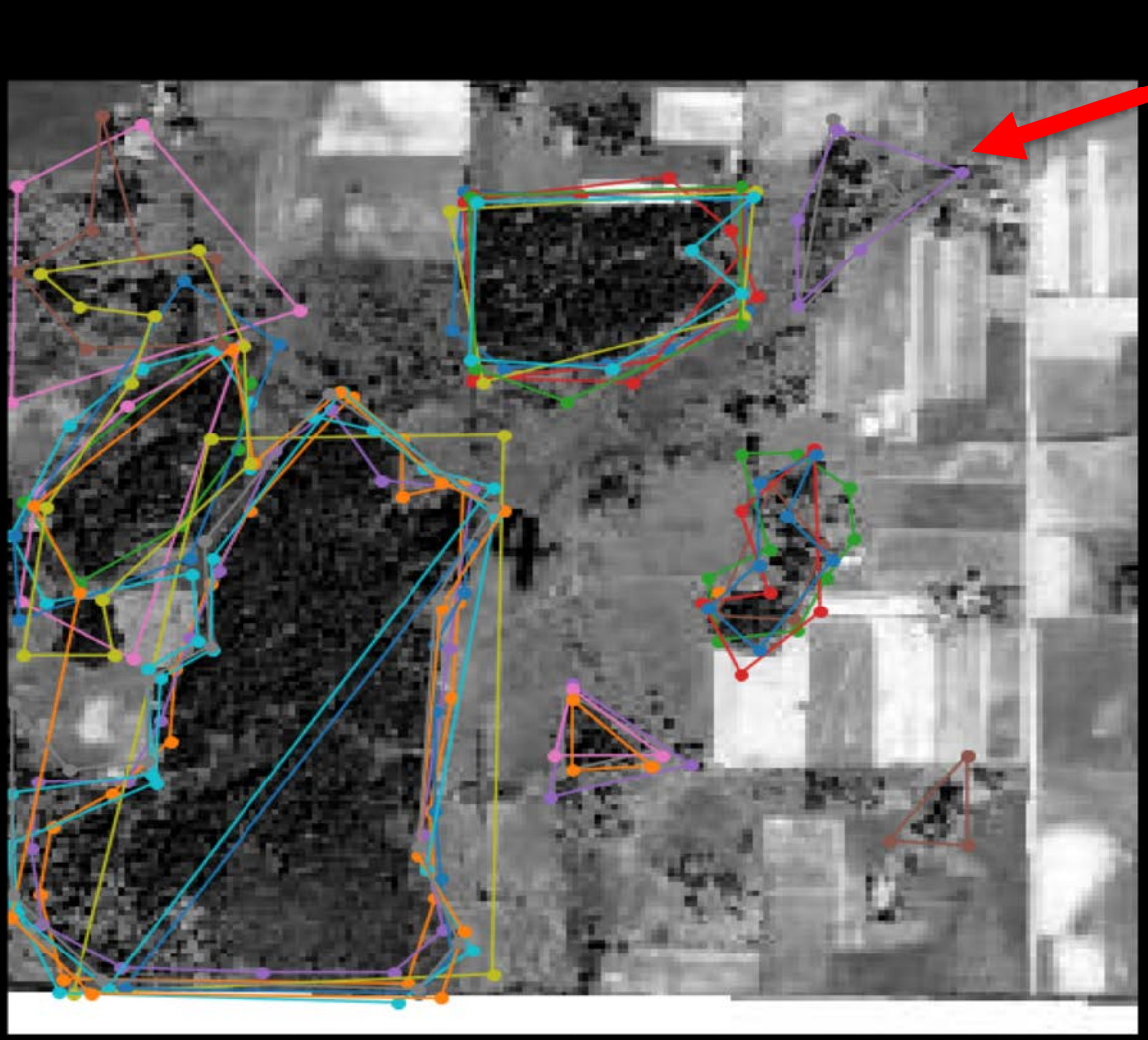


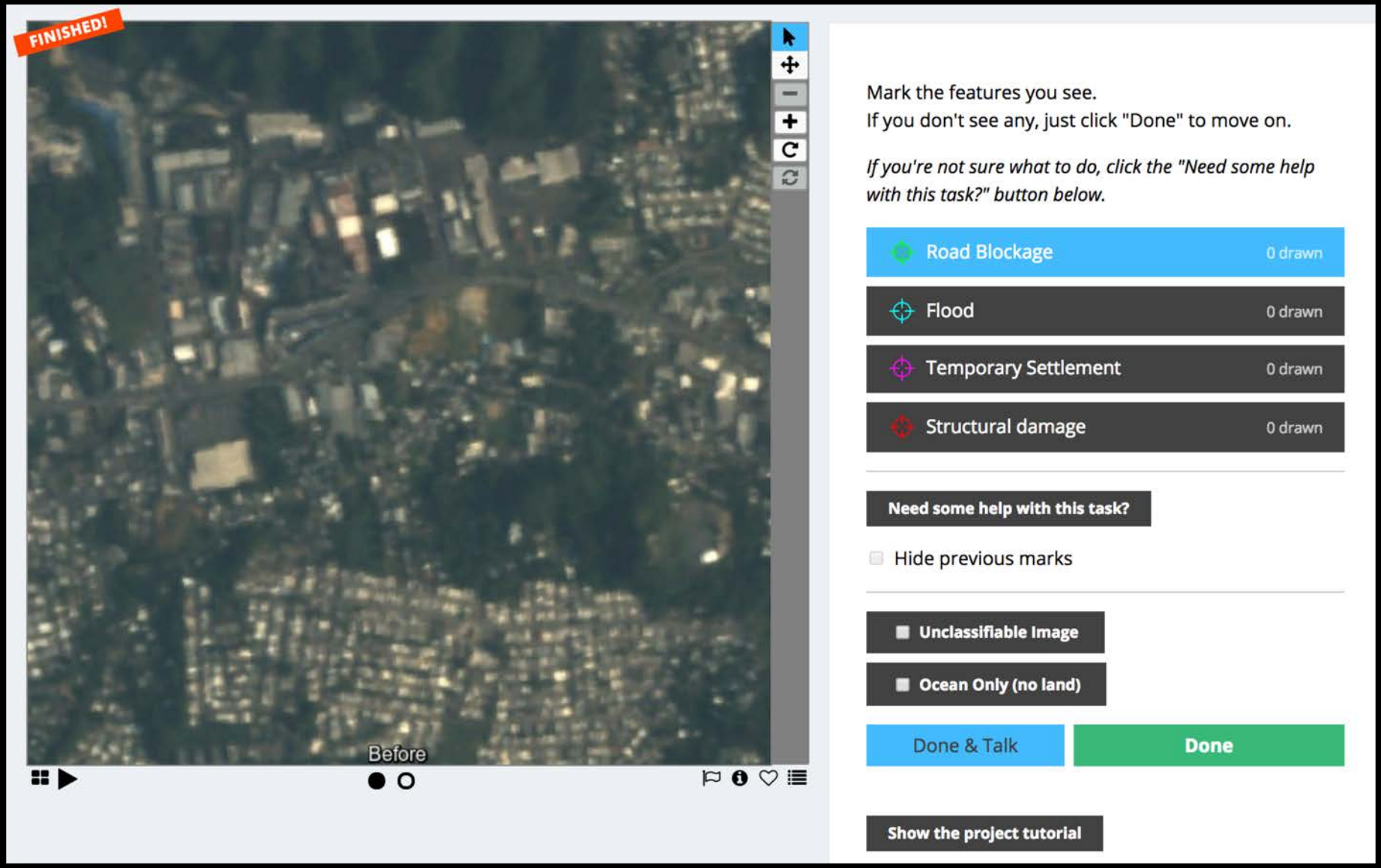
Humans 1-7: Borgholm 1939 image 3



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