

Measuring the ANZACs

Crowdsourcing a complete transcription of WWI soldiers through Zooniverse

*Evan Roberts, Minnesota Population Center
University of Minnesota*

Joint research with
Kris Inwood, Economics and History,
University of Guelph

Les Oxley
Economics
University of Waikato

Collaborating institutions
Archives New Zealand
Auckland War Memorial Museum

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](#)

<http://www.measuringtheanzacs.org/>

Background and motivation: Using
large databases of heights to
measure well-being

Robert William Fogel
NOBEL LAUREATE IN ECONOMIC SCIENCES
and **Stanley L. Engerman**

TIME ON THE CROSS

The Economics of American Negro Slavery



Anthropometric history

- Controversial 1974 book claimed material well-being of slaves similar to free people
- How do you measure material well-being without information on earnings?

Stature as a measure of living standards

- Changes in average stature of a population reflects net nutrition during growth period (0-20 years)
- Reductions (improvements) in net nutrition can come from
 - Lower (greater) food consumption
e.g. price of food rises (falls)
 - Increased (reduced) physical exertion
e.g. Children working more (less) on farms during adolescence
 - Increased (reduced) incidence of disease
(particularly infectious disease)

Nature of Research on Stature

- Height ‘potential’ of most large populations is comparable
- Need individual-level evidence but in the end address the dominant tendencies of a population
- Organize by birth cohorts
- Large samples needed because genetic variation is strong but largely uninteresting.
- Also some effects stronger than others: eg cohort change is glacial but class differences huge implying methodological challenge
- Rely heavily on military and prison sources

Key findings in anthropometric history

- “Antebellum puzzle”: Despite rising incomes, the average stature of American men shrank between 1820 and 1840
- “Industrialization is hazardous to your health”: Despite rising incomes average stature stagnated in 1880s and 1890s in many European countries.
- “Peculiar population”: American slaves severely stunted as children, but stature recovered in adolescence and adulthood
- “Tallest in the world”: Plains Indians were very tall despite population decline

Background and motivation

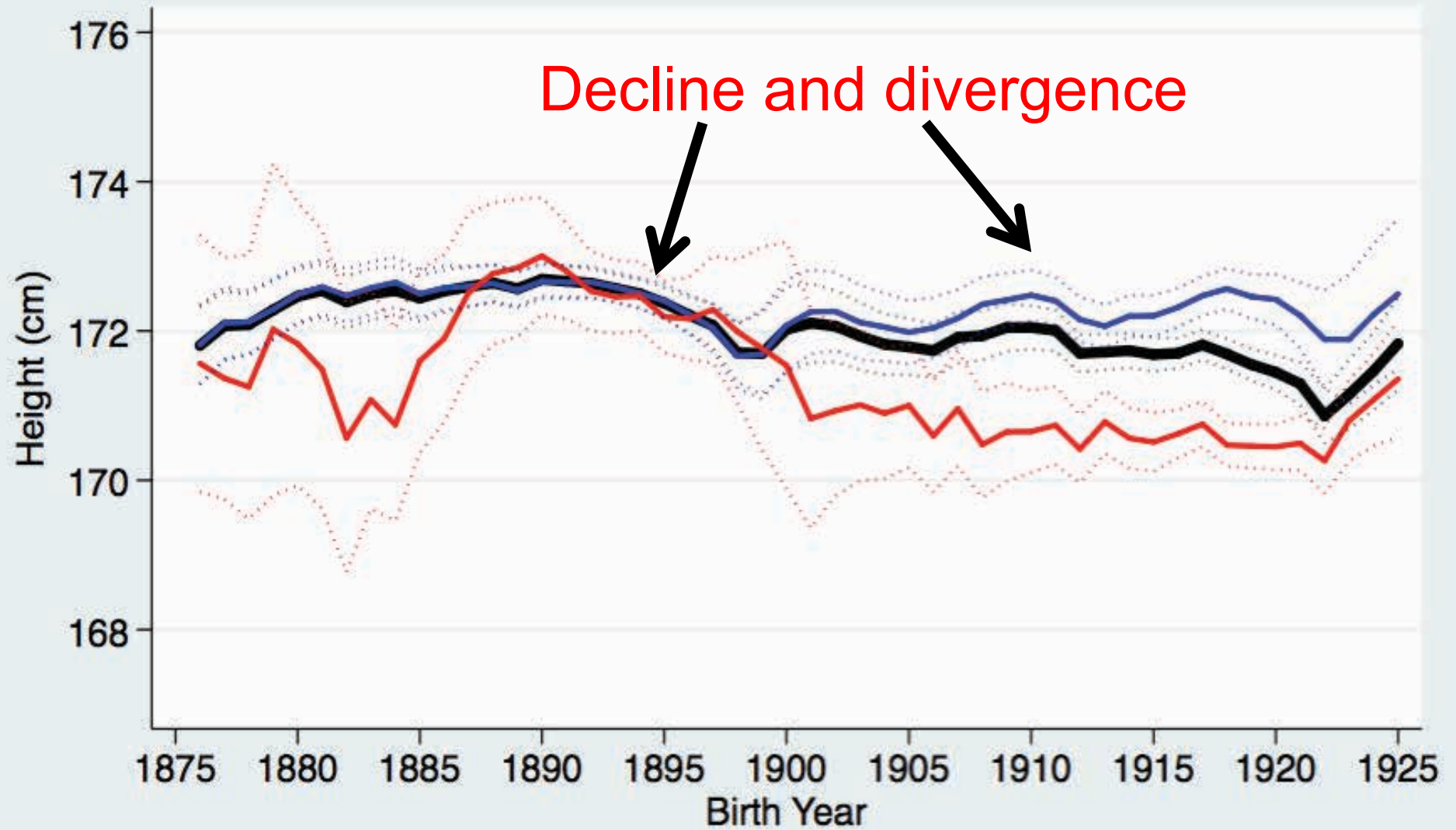
NZ anthropometric history

Our long-term research agenda

- How have health and biological standards of living changed in New Zealand from c.1850 to the present?
- How have differences in health between the indigenous Maori population and European settlers changed from c. 1850 to the present?
- How do biological measures of living standards confirm or refute measures of New Zealand economic growth based on income data?

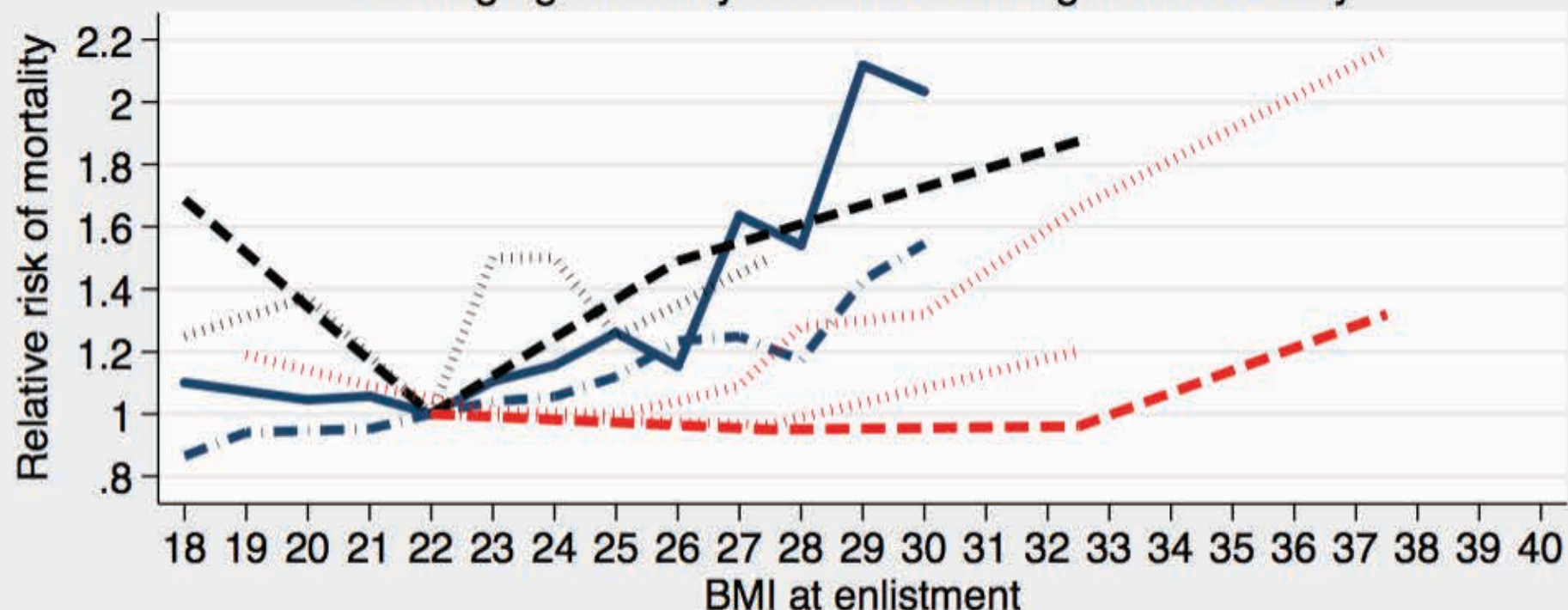
3. Trends in stature in New Zealand

Mean stature of New Zealand soldiers in WWI and WWI



— Average height (cm), overall — Average height (cm), Pakeha
— Average height (cm), Maori

Changing mortality risks of overweight and obesity



- World War I (NZ returned soldiers)
- - - World War II (NZ returned soldiers)
- Murray, 19th century Amherst college students
- - - Costa, 19th Century U.S. Union Army study
- - - Flegal, J. American Medical Association, 2013
- Diverse Populations Collaboration meta-analysis, 2004
- Calle et al, 1999 study of > 1 million U.S. adults

What we've found so far

- Health disparities for Māori emerged several generations after colonial settlement
- Similar findings to US historical studies that mortality risk rises for BMI above 28 (higher end of overweight and obese)
- New Zealand data for World War II suggest a significant reduction in mortality risk between cohorts just 25 years apart. Being overweight is less risky than it used to be [at least for mortality]

Limitations of our current sample

- ~20,000 records [reasonable size in this literature]
- Can't do too many interactions in our models
 - Did Māori do better or worse in particular birth cohorts?
 - Did class differentials vary across time or ethnicity?
- We did not collect combat exposure information
 - (service history sheets)
- Nearly a thousand women served [as nurses]
 - It would be great to include them

[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!](#)

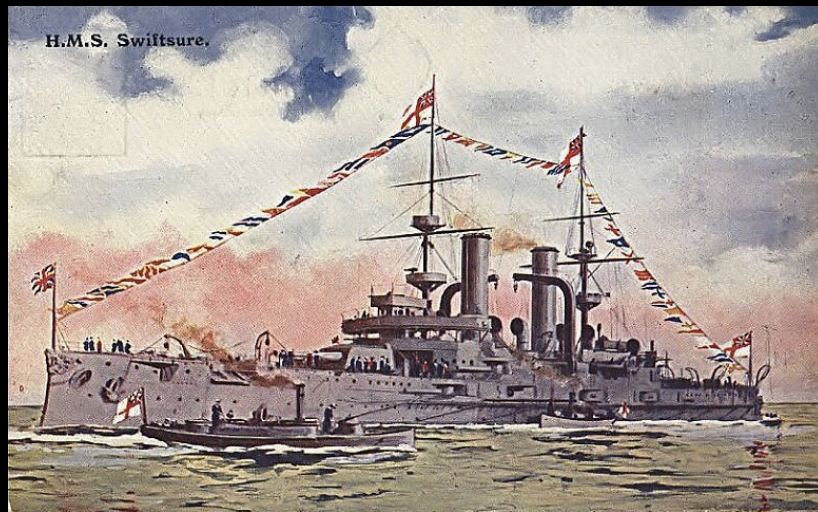
What are these?



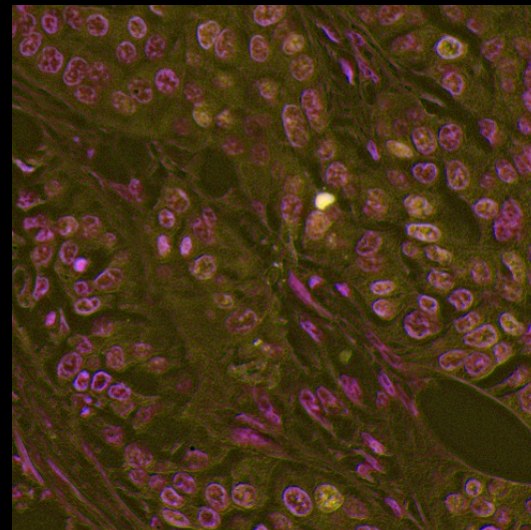
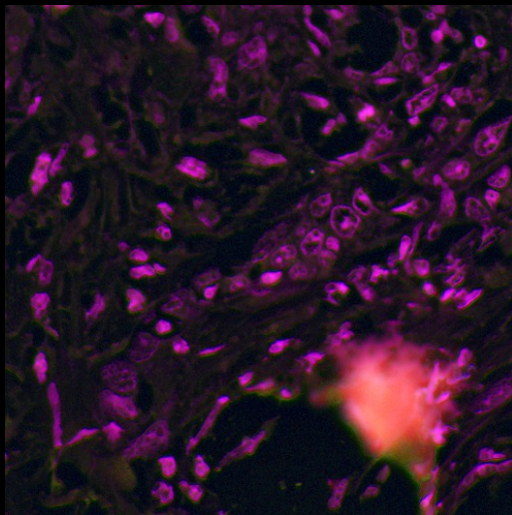
And what do they have to do with these?



Or in fact these?



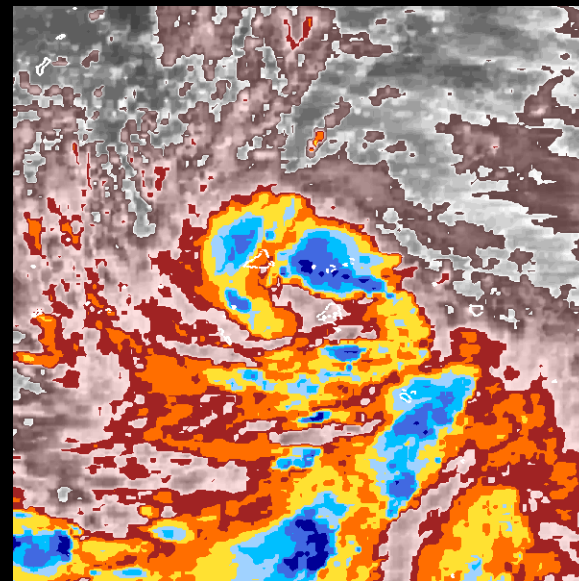
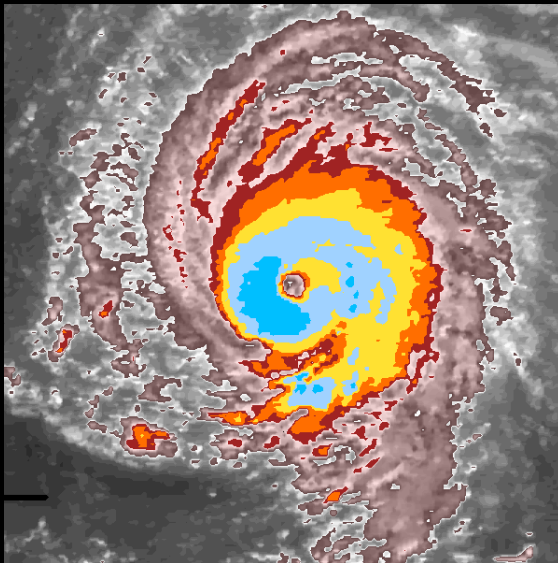
Or these?



Or even these?



Or these?



**These are all images from
research projects that rely on
analysis of complex images
and pattern recognition capabilities
that are currently beyond
computer algorithms!**

Which 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!

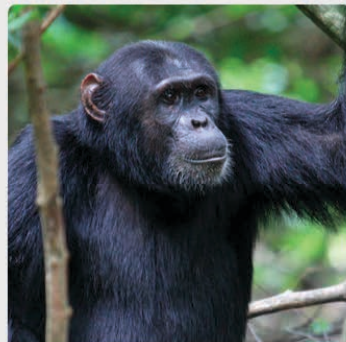
[Get Started](#)



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](#)



Wilbeest Watch

Explore collective intelligence in wildebeest!

[Get Started](#)

Classifying: another human evolutionary tic

Amazingly, only two basic galaxy shapes – but very complex, no two are the same.



Spirals

Lots of star formation so mostly (but not all!) blue-ish.



Ellipticals

Older star populations so mostly (but not all!) red-ish.

**Historically, astronomers would classify galaxies
“by eye” even when 10,000 images!
But what about a million galaxies??**

Astronomy as example of data flood



2.0 meter Sloan Digital Sky Survey Telescope
Apache Point, New Mexico

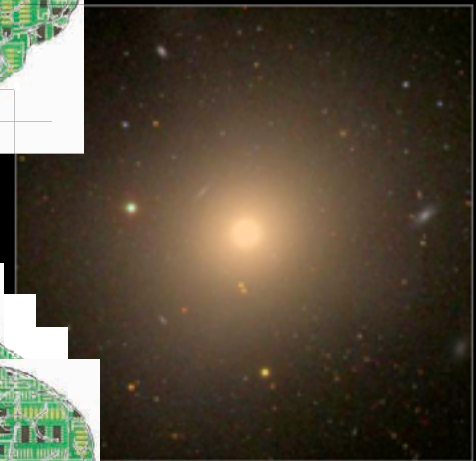
1980 – Palomar Sky Survey:
10,000 galaxies
One expert can classify galaxies
visually.

2000s - Sloan Digital Sky Survey:

~1000000 galaxies

**Need computers
to process that
much data!**

But how can we “computerize”
the best algorithm that can
process complex data?



Galaxy Zoo launched in 2007 inviting public to classify galaxies



In 1.5 years, 35 million classifications by ~150,000 users

Roughly 3.3 continuous person-years!

Tutorial: Simple, divisible tasks

Part 1B ... More Tricky Spiral or Elliptical Galaxies

Some galaxies are a bit more tricky. As you noticed in the previous section, some spiral galaxies can look like ellipticals when viewed edge-on. Also, in some faint spiral galaxies, you have to look hard to see the spiral arms. Now, see if you can separate the genuine ellipticals from the spirals.

Try your hands at some!

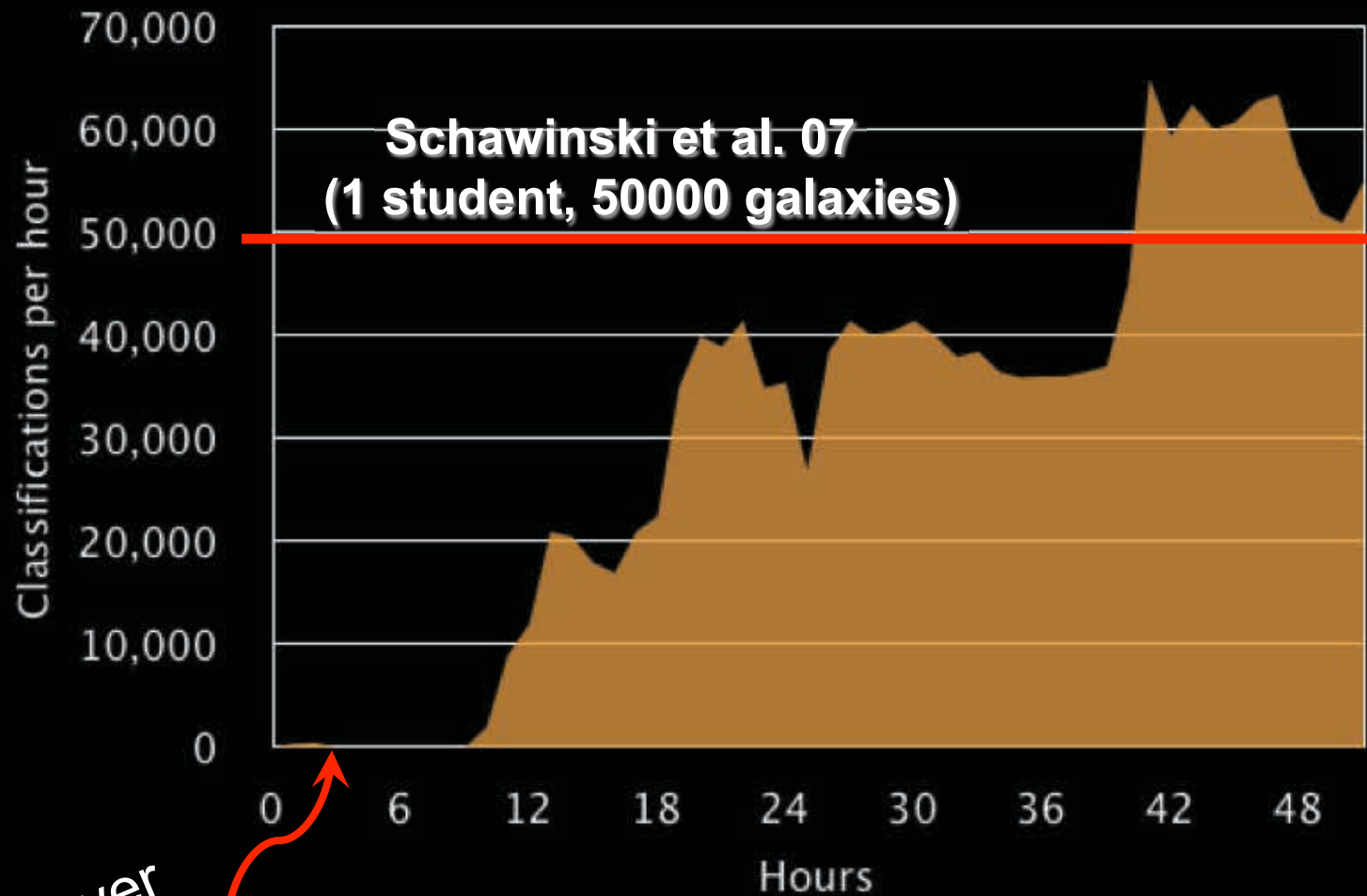
Click the image to see if you're right.



Part 1C ... Merging Galaxies

Sometimes, galaxies crash into each other, or come close. These are called merging galaxies. Merging galaxies are very interesting to astronomers because we think that large galaxies are built from mergers of small galaxies – if we see merging galaxies, we can see a snapshot of how that process happens. When you look for mergers, look for places where two galaxies appear to be merging into one. The galaxies should be close together, and you should be able to see some connection between them. In the trial or in your galaxy analysis, whenever you see this, click the button that says "Mergers".

The first two days...



Server
fried at 3
hours in!

GALAXY ZOO.org

Welcome

Home

The Science

How to Take Part

Galaxy Analysis

Forum

Press & News

FAQ

MOST POPULAR STORIES NOW

MOST E-MAILED

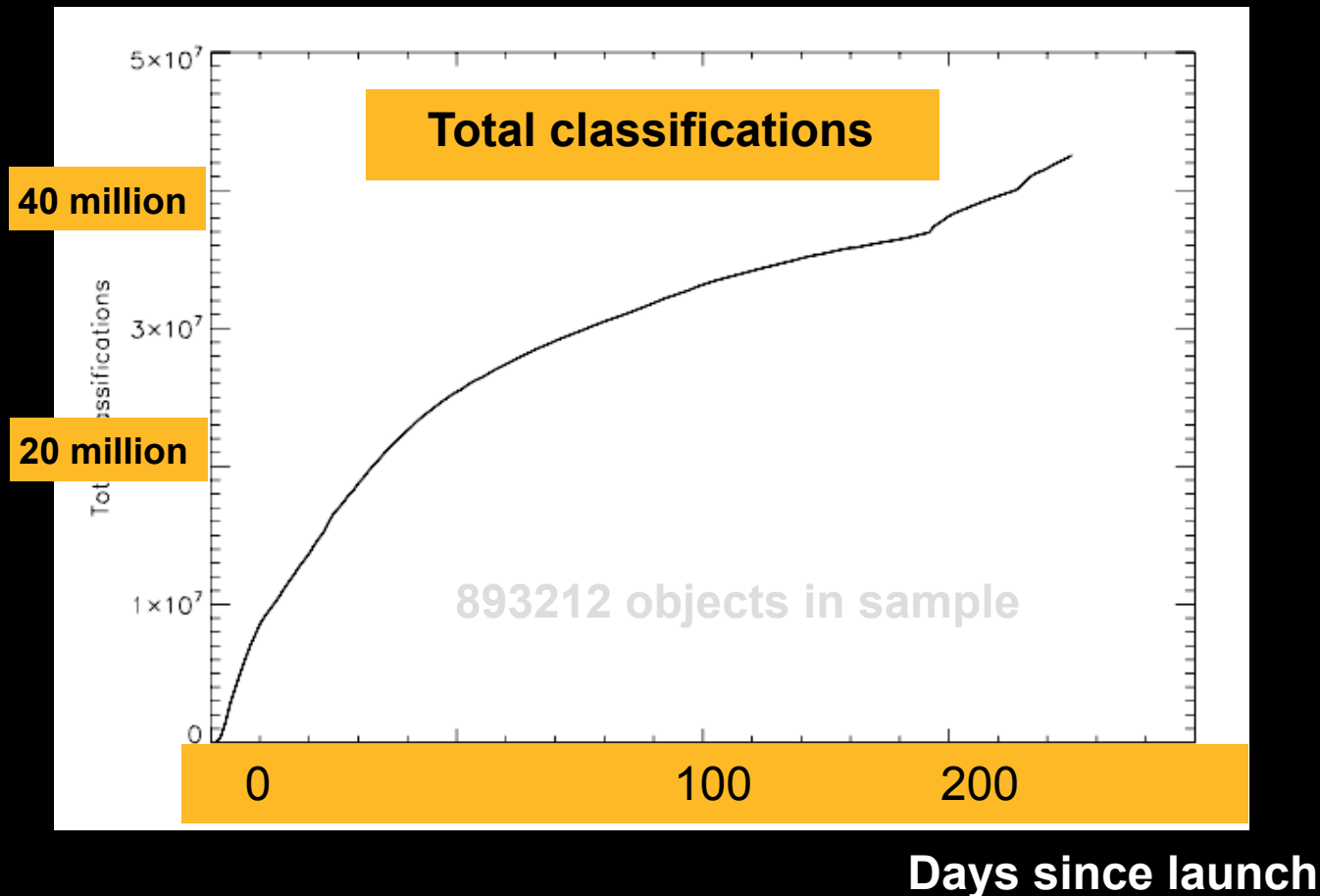
MOST READ

- 1 Moles 'good indicator to ageing'
- 2 Huge dog is reluctant media star
- 3 Scientists seek galaxy hunt help
- 4 'My mother held me down'
- 5 Man flies to wedding a year early

► Most popular now, in detail

Public outreach and media presence
to engage broad and large audience

First six months



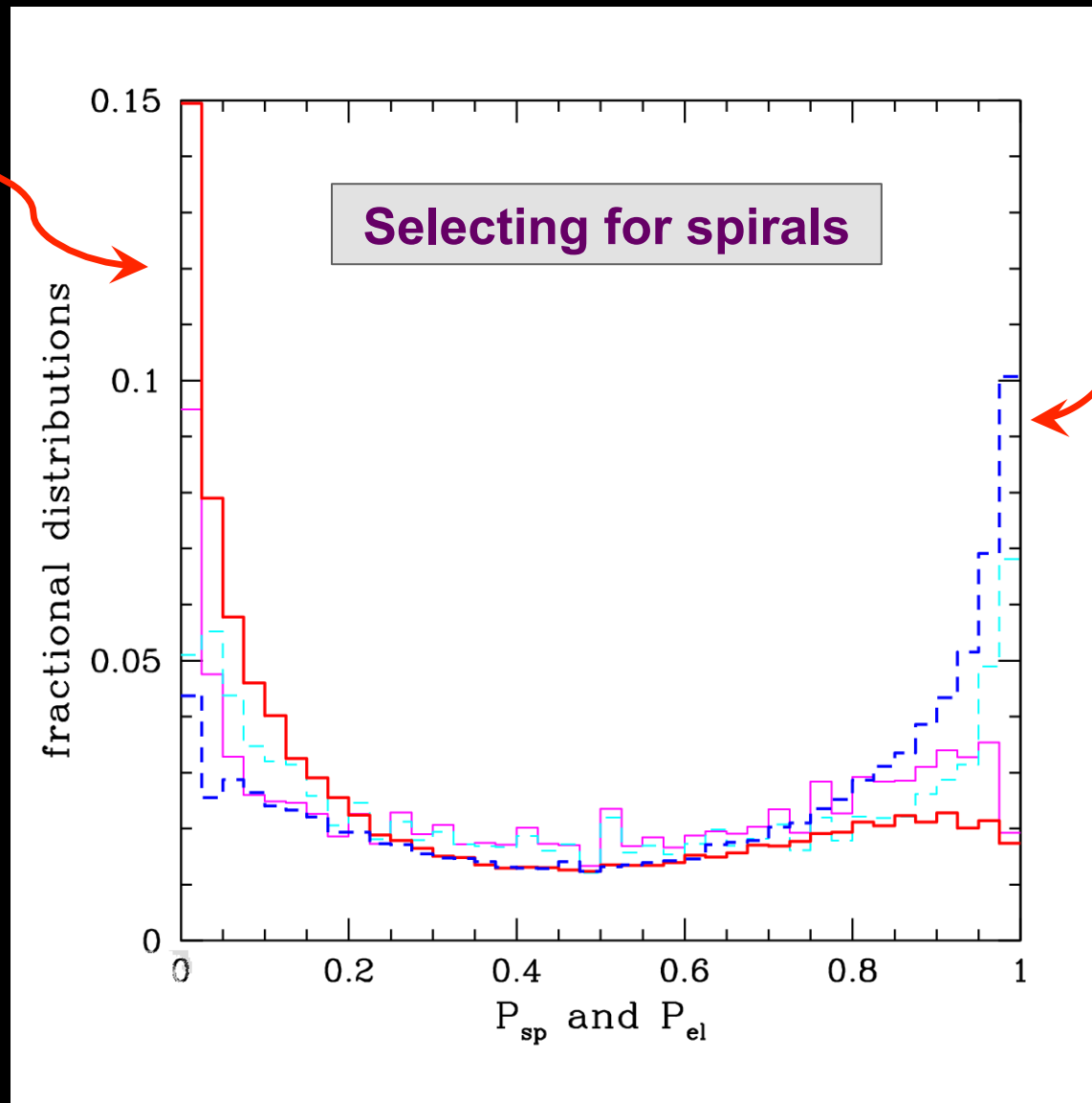
After 'cleaning' raw clicks:

34,617,406 classifications by 82,931 users

Roughly 3.3 continuous person-years!

Combining clicks

“Pure”
elliptical
sample



“Pure”
spiral
sample

Galaxy Classification – now as probabilities!

The case for crowdsourcing your research!

The scale of the problem – what do we do with 40 Tbytes a day?

- Largest professional classification is only ~5% of SDSS
- Galaxy Zoo provided 3.3 person years in first 6 months

(Measurable) Accuracy – wisdom of the crowds

- Multiple independent classifications give us an estimate of error
- For a set of noisy data, enough inexpert classifiers will produce more accurate classifications than an expert classifier (where 'enough' depends on how noisy the data, and how inexpert and expert the classifiers).

OK – BUT!!

1. Can we apply this to research beyond galaxy zoo?
2. Can we get the attention of the crowd?

The Data Deluge



BIG DATA!!

**Not just Astronomy
suffering from Big Data!**



Analysis
Gap

Moving Beyond Galaxy Zoo


- “Information age” = Data-Intensive Research

Researchers overwhelmed with data sets from many different sources

- Data captured by instruments
- Data generated by simulations
- Data generated by sensor networks
- Data generated through digital recovery methods
- eScience is the set of tools and technologies to support data federation and collaboration
 - For analysis and data mining
 - For data visualization and exploration
 - For scholarly communication and dissemination

eScience is also the technological backbone
that enables today's Citizen Science

Moving Beyond Galaxy Zoo

1. **Passive** – distributed computing so no involvement from volunteer other than their computer (SETI@Home, Einstein@Home)
 2. **Data Collection** – distributed sample collection/ observation (**bird counts**, variable star observations, weather data, **community-based documentation**)
 - “Classic” citizen science solving the need for distributed sensors as well as current internet-enabled data collection projects.
 3. **Data Analysis** - distributed data analysis: enabled by the internet (**Zooniverse**, FoldIt!)
 - Solution to large amounts of complex data where volunteers are part of data-processing pipeline for complex systems e.g. pattern matching
- 



Projects

About

Talk

Collect

Daily Zooniverse

Blog

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!](#)



10,000
volunteers

classifications

150,000



- Whale FM: Listen to whale calls to decode their language

whale.fm



Welcome to the Whale Song Project
You can help marine researchers understand what whales are saying. Listen to the large sound and find the small one that matches it best. Click 'Help' below for an interactive guide.

Home About ▾ How to Take Part ▾ Account ▾

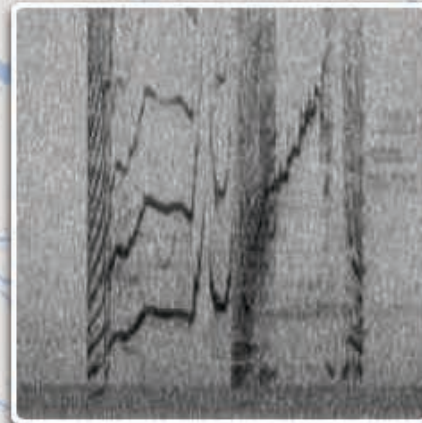
Help



18 MAY 2009 08:54

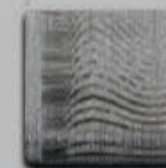
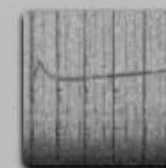
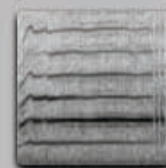
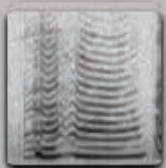
TAG: DELPHINUS 1

TRACKING ☐ OFF



Google

Map data ©2012 GIS Innovatsia, Google - [Terms of Use](#)



Old Weather: Ships logs have a similar data structure and transcription issues to military personnel files

H.M.S. "Pegasus Th", Thursday 2nd day of

From Archangel, To Dundee or At

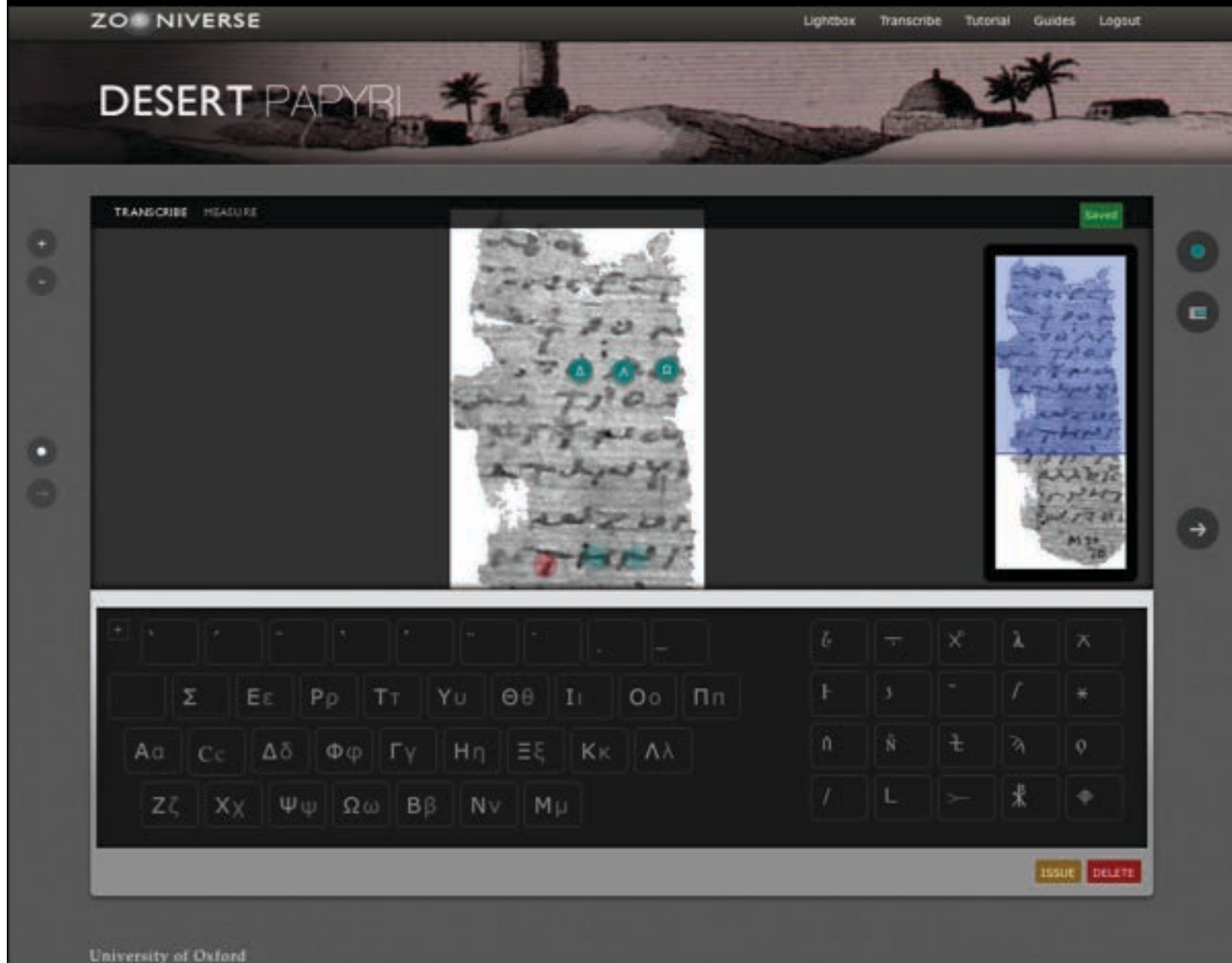
Time	Patent Log	Distance Run		True Course	Revolutions per minute	Wind		Weather	State of the Sea	Height of Barometer and attached Thermometer	Temperature			Position	
		Miles	Tenths			Direction	Force				Air	Wet Bulb	Sea	0800	2000
0100	1399	10	8	357	161										
0200	1410.1	11	1	—	"										2.30 ex. 3e
0300	1422.5	12	4	—	"										
0400	1438.2	10	7	—	"	SxS	8	oca b		29.81 56	52	50	42	14.0	
0500	1445.6	12	4	—	"										6.0 Hands
0600	1456.0	10	4	—	"										
0700	1466.2	10	2	—	"										
0800	1479	12	8	—	"	SxS	8	oca b		29.66 58	52	50	41	8.0	9.0 Sighted Fel. l
0900	1490.3	11	3	—	"										9.15 9e S 25W
1000	1500.9	8	0	—	"										10.22 S 10W 4.

25,000 volunteers
1,000,000 pages



Old Weather: transcribe the logs of WW1 vessels as they travel the globe, collecting valuable climate data

Ancient Lives Project: transcribing 2000 year old papyri fragments from Oxyrhynchus, Egypt. Partnership between UMN and Oxford, UK.

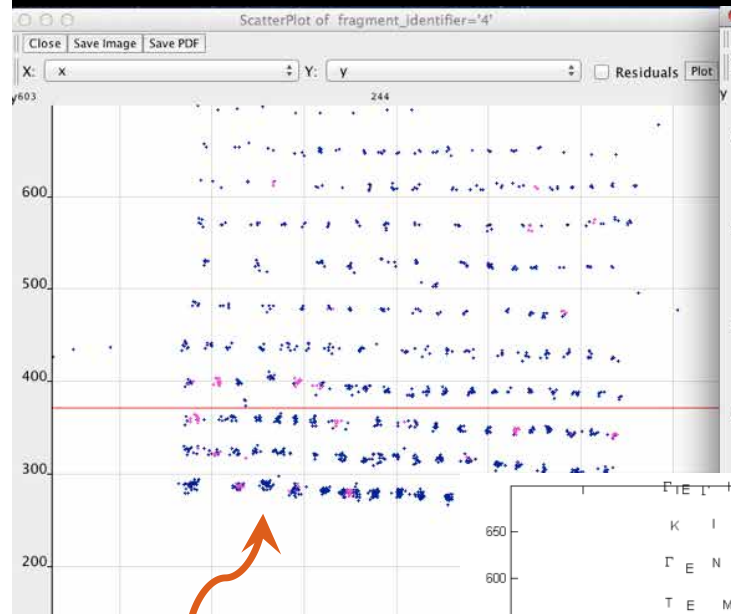


300,000
volunteers

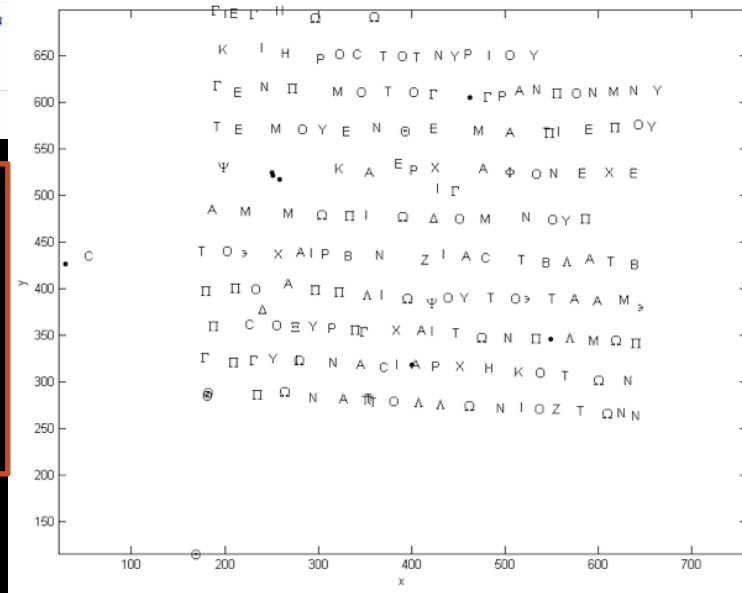
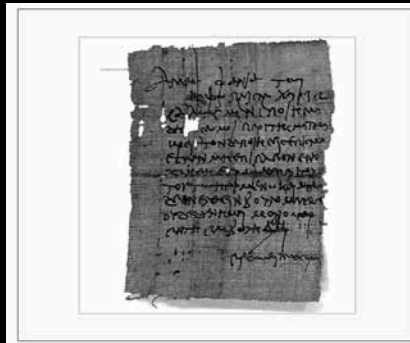
classifications
100m



Ancient Lives Project: UMN developed the pipeline to turn clicks into consensus into transcribed text that can be edited & curated.



Centroiding raw clicks and aligning centroids to factor out scribe's slant.



Identifying most probable letter with each centroid.

1 Annotations

2 Text Curation

3 Ready to Submit

[Click to show original Generated text](#)

Curated text

1	A	MΘP	ΦANA	T
2	Φ	ΠM	XΠ	
3	AΛΣ	CΩ	NI	O ^{EE} M
4	Δ	Ω ^{EE} Π	PITECΣ	ON
5	ΘP	O NΔΠΑH	I	
6	E	MHENI	ΠΛON	O
7	XΛ			
8	TO		NΩ	
9		NBOT		
10		EO		
11	NT	B		

Line

5

Column

7

Current

O

Action

New Content

- One or more characters for "insert character(s)"
- One character for "replace"
- A line of text for "insert line"

Annotation and curation tool displaying original image and consensus transcription that can be edited or augmented with metadata.

Anatomy of a Zooniverse Project

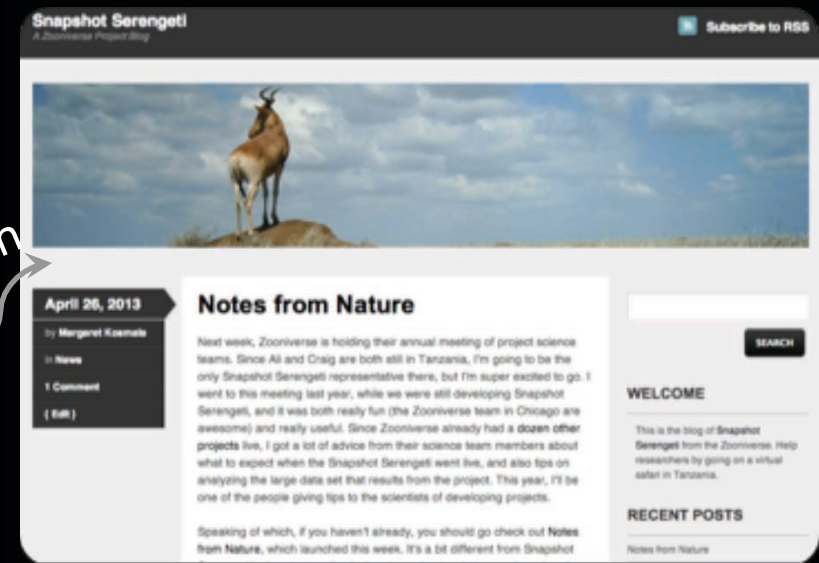
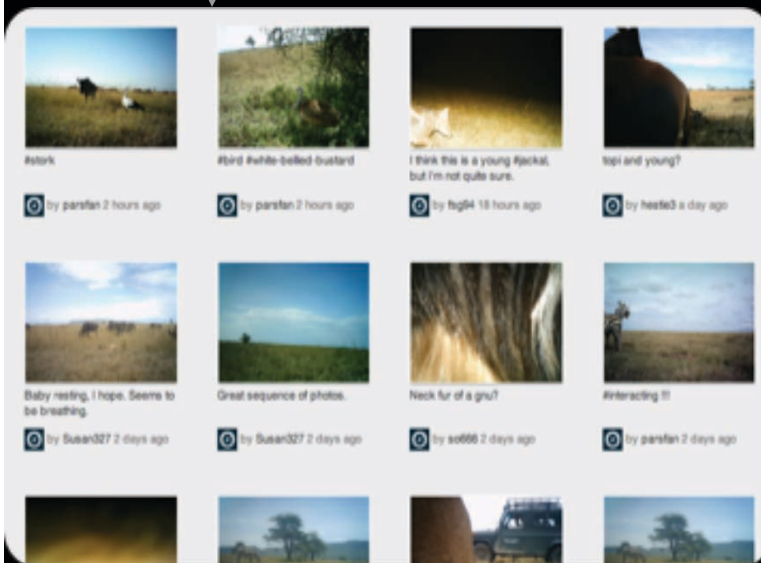


“Talk” – where volunteers discuss their findings and ask for help.

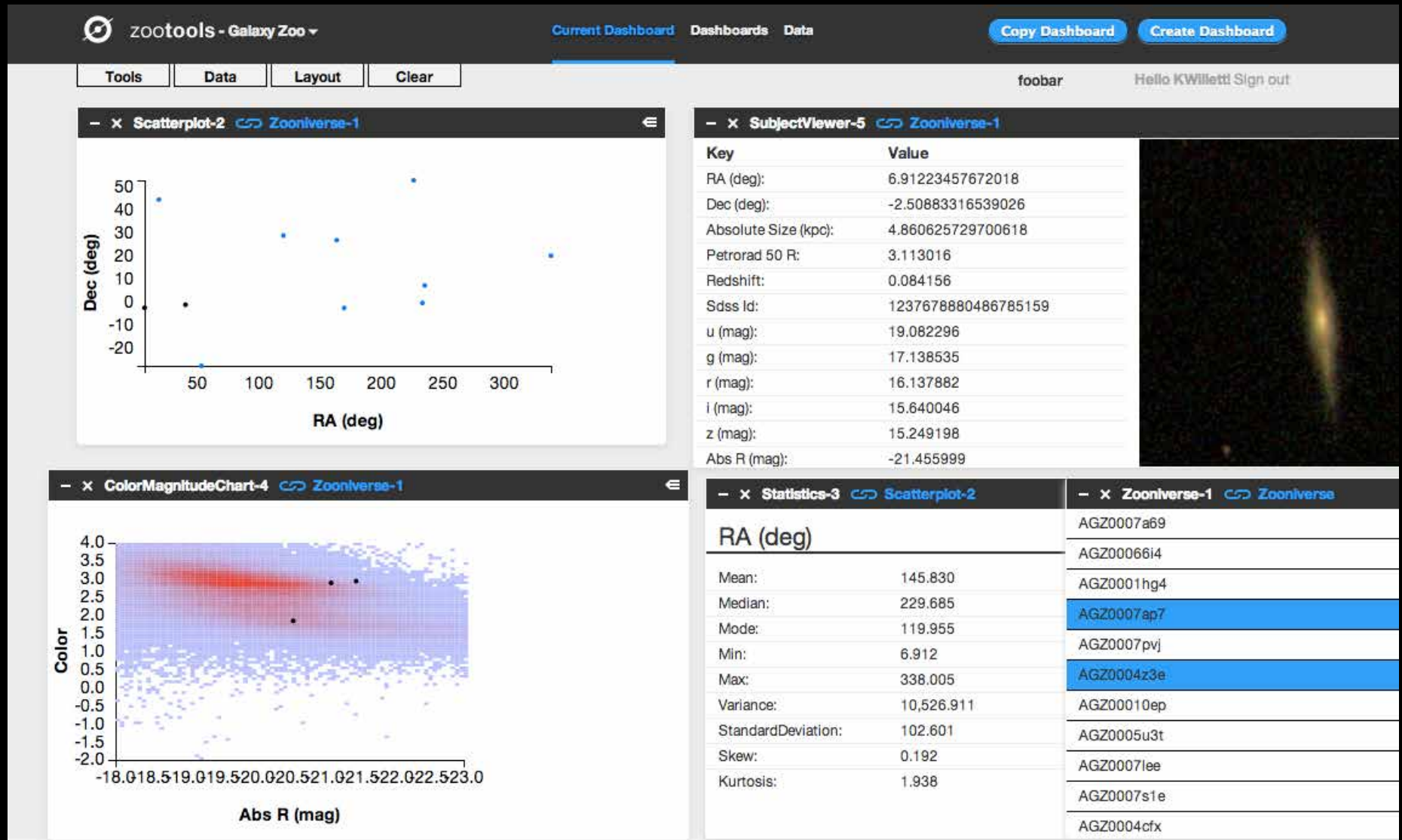


Primary task interface

“Blogs” – where volunteers hear from the research team.

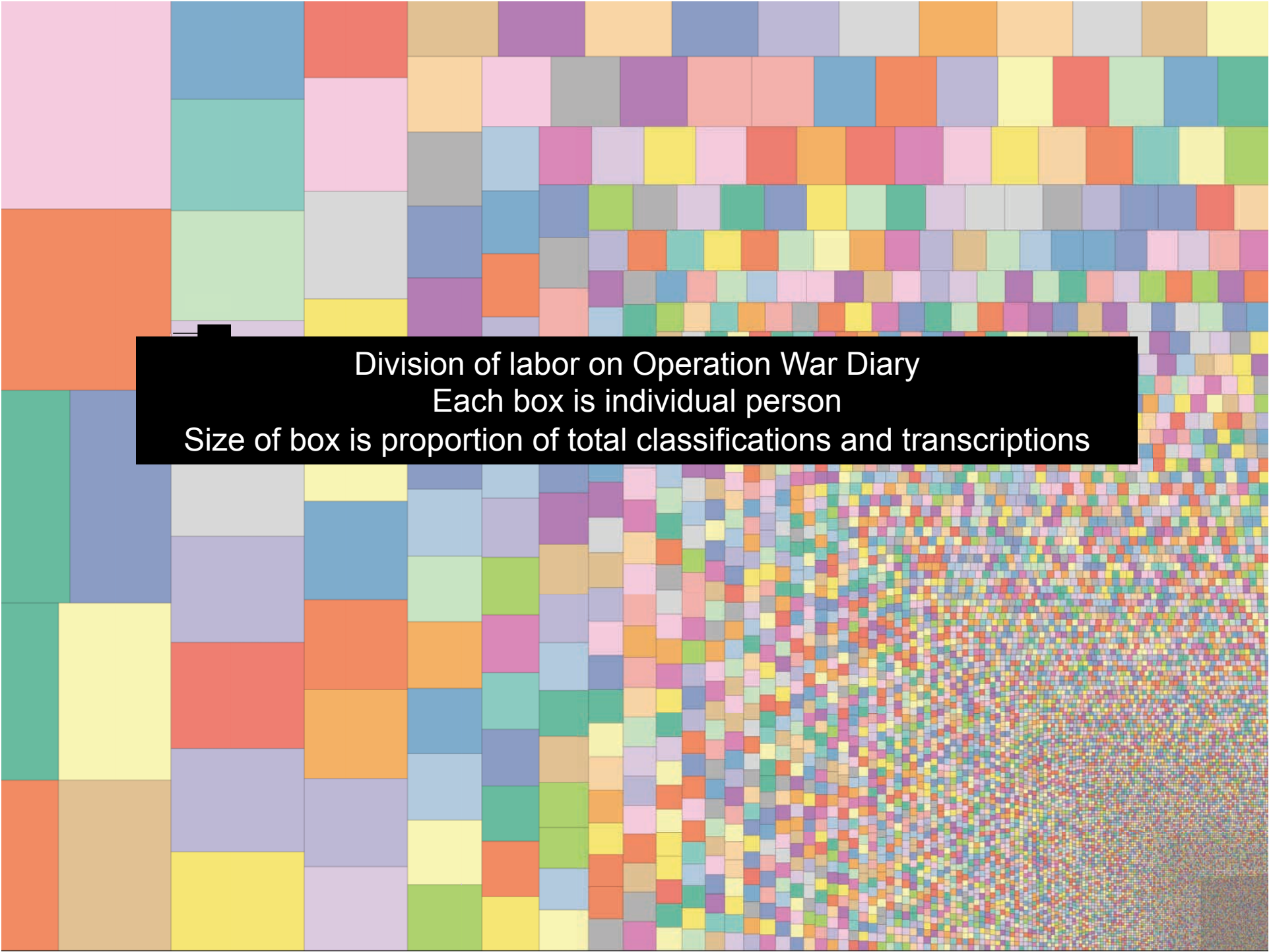


Anatomy of a Zooniverse Project



Data analysis tools - enabling volunteers to explore the data:

tools.zooniverse.org



Division of labor on Operation War Diary
Each box is individual person
Size of box is proportion of total classifications and transcriptions

This treemap visualization represents the division of labor on Operation War Diary. The entire image is composed of numerous colored rectangles of varying sizes, where each rectangle represents an individual person's contribution. The size of each box is proportional to the number of classifications and transcriptions that person completed. The colors of the boxes are used to categorize the individuals, with a variety of colors including shades of pink, blue, orange, green, purple, and grey. The layout is hierarchical, with larger boxes on the left and top-left representing more significant contributors, and smaller boxes becoming increasingly dense towards the right and bottom-right, illustrating the vast scale of the project.

Most of the work being done by relatively
small number of citizen scientists!

How do their insights and knowledge
reach the research teams?

Classes of Knowledge Discovery



“There are known knowns. These are things we know that we know. There are known unknowns. That is to say, there are things that we know we don't know. But there are also unknown unknowns. There are things we don't know we don't know.”

Donald Rumsfeld

Classes of Knowledge Discovery:

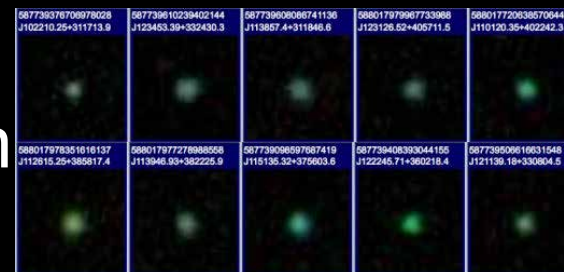
Known knowns : Primary task. Data reduction by science team.



Known unknowns : Related to primary task. Results funneled to specific researchers.



Unknown unknowns : Serendipity. Currently rely on forum moderators to filter.



Accelerating Knowledge Discovery:

- Old-fashioned “linear” forum tool couldn’t support 600,000 people over multiple projects with different levels of engagement discussing their finds
 - Scientists could not easily find serendipitous discoveries posted on the forum
- The social media (communication activities) not totally integrated with data filtering activities.
 - Lack of “easy-transfer” between primary task and forum ended up losing people on both sides

Want to make the most out of user’s desire to conduct their own research and accelerate knowledge discovery.

Key is in communication with and between volunteers:

- Develop “object oriented” discussion tool linked directly from classification task

[“talk.galaxyzoo.org”](http://talk.galaxyzoo.org)

Each project has own “talk” URL

- More engaging: integrates social media aspects including tagging, creating and sharing collections, discussing observations with peers and team researchers

- Develop a Citizen Science Journal allowing users to publish their work

[“letters.zooniverse.org”](http://letters.zooniverse.org)

- Develop more advanced volunteer-based research enabling tools

[“tools.zooniverse.org”](http://tools.zooniverse.org)

Beta version!

Beta version!

Talking in the Zooniverse



VISIT MEASURING THE ANZACS

CLASSIFY

TALK

Measuring the ANZACs Talk

Search...



Notes

General comment threads about individual subjects

 [Evan Roberts, Measuring the ANZACs](#) Field Guide [an hour ago](#)


 10 Participants

 25 Discussions

 61 Comments

Troubleshooting

Having technical issues with Measuring the ANZACs in particular? Let us know about them here!

 [Bonnie123](#) Boxes disappearing? [16 hours ago](#)

 24 Participants

 15 Discussions

 83 Comments

The Recruiting Station

Just arrived? Say hello here.

 [Evan Roberts, Measuring the ANZACs](#) Welcome to Measuring the ANZACs [3 days ago](#)

 10 Participants

 3 Discussions

 21 Comments

Measuring the ANZACs Discussion

Discussion about the Measuring the ANZACs project.

 [Evan Roberts, Measuring the ANZACs](#) Magnification [5 days ago](#)

 4 Participants

 1 Discussion

 4 Comments

“Talk” entered directly through “talk” website shows boards, collections...

Talking in the Zooniverse

Measuring the ANZACs Talk

[Measuring the ANZACs Talk](#) > Notes

Search...



Notes

+New Discussion

Field Guide

 [Evan Roberts, Measuring the ANZACs](#) *an hour ago* A lot of folks have asked about the different types of documents. We have a s...

1 Participant

1 Comment

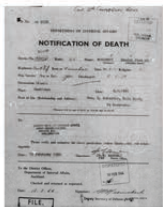


Subject 1016353

 [Evan Roberts, Measuring the ANZACs](#) *18 hours ago* hi @Zanna640. The keyword here is Ballot as you've recognized, so that would ...

2 Participants

2 Comments



Subject 994216

 [Bonnie123](#) *a day ago* Thanks Evan for the response. Also, Thanks so Much to all the individuals on...

4 Participants

7 Comments

Citizen scientists have asked a lot of questions about the “objects” (pages in personnel files). We’ve been responding!

“Talk” entered through classification interface provides “microblog” or “tweet”-like comment box.

What Zooniverse does and will do

- Large volume of complex images – big data challenge
 - Crowdsourcing - Human processing in combination with machines
 - is one solution
- Zooniverse: largest crowdsourcing research platform
 - At the same time we do research on the science of crowdsourcing (how to optimize human-machine partnership -> optimizing human attention and cycle of machine learning)
 - Provides education opportunity for millions of members of the public as well as in formal education
- “Talk”: tool to engage volunteers in collaborative discovery
- **Future: the Zooniverse has developed a “Do it Yourself” API that will allow anyone with a good chunk of data to put forth a Zooniverse project as easy as putting out a blog [on a webserver you manage yourself].**

<https://github.com/zooniverse/scribeAPI>

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](#)

What Measuring the ANZACs will do!

- Partial transcription of all 140,000 files
 - Millions of images!
 - Multiple pages and forms per file
- Classify page types
- Capture data from attestations (enlistment), service histories, and death notifications
- Timely public history engagement with community
 - Aim: Finish by end of the war's centenary

NEW ZEALAND EXPEDITIONARY FORCE.

ATTESTATION OF

No. 1310 Name: Smith Robert John Regiment or Unit: H Coy 11th

Questions to be put to the recruit before enlistment.

1. What is your name? ... 1. Robert John Smith
2. Where were you born? ... 2. London
3. Are you a British subject? ... 3. Yes
4. What is the date of your birth? ... 4. 1875
5. What is your trade or calling? ... 5. Painter
6. Are you an indentured apprentice? If so, state name and to whom? ... 6. No
7. What was the address at which you last resided? ... 7. 11th Avenue
8. Have you passed the Fourth Educational Standard or its equivalent? ... 8. Yes
9. What is the name and address of your present or last employer? ... 9. Chambers's Patent
10. Are you married? ... 10. No
11. Have you ever been sentenced to imprisonment by the Civil power? If so, when and where? ... 11. No
12. Do you now belong to any military or naval force? If so, to what corps? ... 12. No
13. Have you ever served in any military or naval force? If so, state which and cause of discharge. ... 13. Yes
14. Have you truly stated the whole (if any) of your previous service? ... 14. Yes
15. Have you been registered for compulsory military training under the Defence Act, 1906? If so, where? ... 15. No
16. Have you ever been rejected as unfit for the military or naval forces of the Crown? If so, on what grounds? ... 16. No
17. Are you willing to be vaccinated or revaccinated and inoculated? ... 17. Yes
18. Are you willing to serve in the Expeditionary Force in or beyond the Dominion of New Zealand under the following conditions, provided your services should so long be required: For the term of the present European war and for such further period as is necessary to bring the Expeditionary Force back to New Zealand and to disband it? ... 18. Yes

Note.—Your discharge will not be granted before your return to New Zealand unless permission for discharge elsewhere be obtained from the G.O.C. the New Zealand Expeditionary Force.

I, Robert John Smith, do solemnly declare that the above answers made by me to the above questions are true, and that I am willing to fulfil the engagement made.

Signature of Recruit: Robert John Smith

Signature of Witness: W. J. ...

Oath to be taken by recruit on attestation.

I, Robert John Smith, do solemnly promise and swear that I will be faithful and bear true allegiance to our Sovereign Lord the King, his Heirs and Successors, and that I will faithfully serve in the New Zealand Military Forces, according to my liability under the Defence Act, and that I will observe and obey all orders of His Majesty, his Heirs and Successors, and of the Generals and Officers set over me, until I shall be lawfully discharged. So help me, God.

Certificate of Magistrate or Attesting Officer.

The above questions were read to the above-named recruit in my presence. I have taken care that he understands each question, and that his answer to each question has been duly entered as replied to, and the said recruit has made and signed the declaration and taken the oath before me, at London, N.Z., on this ... day of ...

191

Signature of Attesting Officer: ...

If any alteration is required in this page of the Attestation, the Attesting Officer should be requested to make it and initial the alteration.

Who are you and where are you from?

What's your job?

How can we contact Mum & Dad or your Missus when you die?

Ever served before?

Do you swear to obey orders?

Reinforcement: 12 **HISTORY - SHEET** Service despatched to soldier: 4-3-21 ✓
 Date of birth: 18/01/1880 Rank: Pte Surname: Smith Christian Name: Robert John No. 13130

Occupation: Contractor Religion: Ireland Last New Zealand address: Te Wharua
 Last employer: Mr. Rimmer C of C Hastings

Name, relationship, and address of next-of-kin (if not resident in New Zealand, insert also name and address of nearest relative in New Zealand):
 Mr. Violet Smith 5 Railway Terrace Hastings
 Mr. L. Graham 35 Marine Parade Hastings
 Mr. Rimmer Te Wharua

Service
 No. 13130 Rank: Pte
 Name: Robert John Smith
 Address: 35 Railway Terrace Hastings

Periods of Service. Theatres of Operation.
 In New Zealand: 132 days. Australasian
 Overseas: 3 years 153 days. Egyptian
 Total service: 2 years 275 days. Egyptian E.F.
 Date commenced duty: 4. 1. 16. Balkan
 Date finally discharged: 5. 10. 18. Western 1916-17.
 European
 Asiatic

Wounds
 Decorations: MEDAL OF ACTION
 BRITISH WAR MEDAL
 VICTORY MEDAL
 The foregoing particulars are correct.
 Signature: J. J. Smith
 N.B. Do not omit to advise this office of any future change of address.

Sick
 Total service: 2 years 275 days. Egyptian E.F.
 Date commenced duty: 4. 1. 16. Balkan
 Date finally discharged: 5. 10. 18. Western 1916-17.
 European
 Asiatic

Killed in action
 Died of wounds
 Missing
 Prisoner
 Injuries in or by
 Service
 Discharge

Decorations: BRITISH WAR MEDAL 18 MAY 1919
 VICTORY MEDAL 18 MAY 1919
 The foregoing particulars are correct.
 Signature: J. J. Smith
 N.B. Do not omit to advise this office of any future change of address.

Service (Where did he go? What did he do?)

Wounds

Sickness

Death

Prisoner

Discharge

Measuring the ANZACs data structure

- 147 variables/fields/columns from source material
 - Some variables describe categories of events that can occur multiple times
 - Sick, wounded, transfers between units, entering new combat zone
 - You only die once
- 6 different document types to identify
- Paradata : Data about the transcription of each element
 - 17 elements of paradata about each transcription or classification! Includes who, when, X/Y co-ordinates on page
- Significant automatically generated metadata

No. 814234

1	338
---	-----

* Strike out words not required.

if you have previously

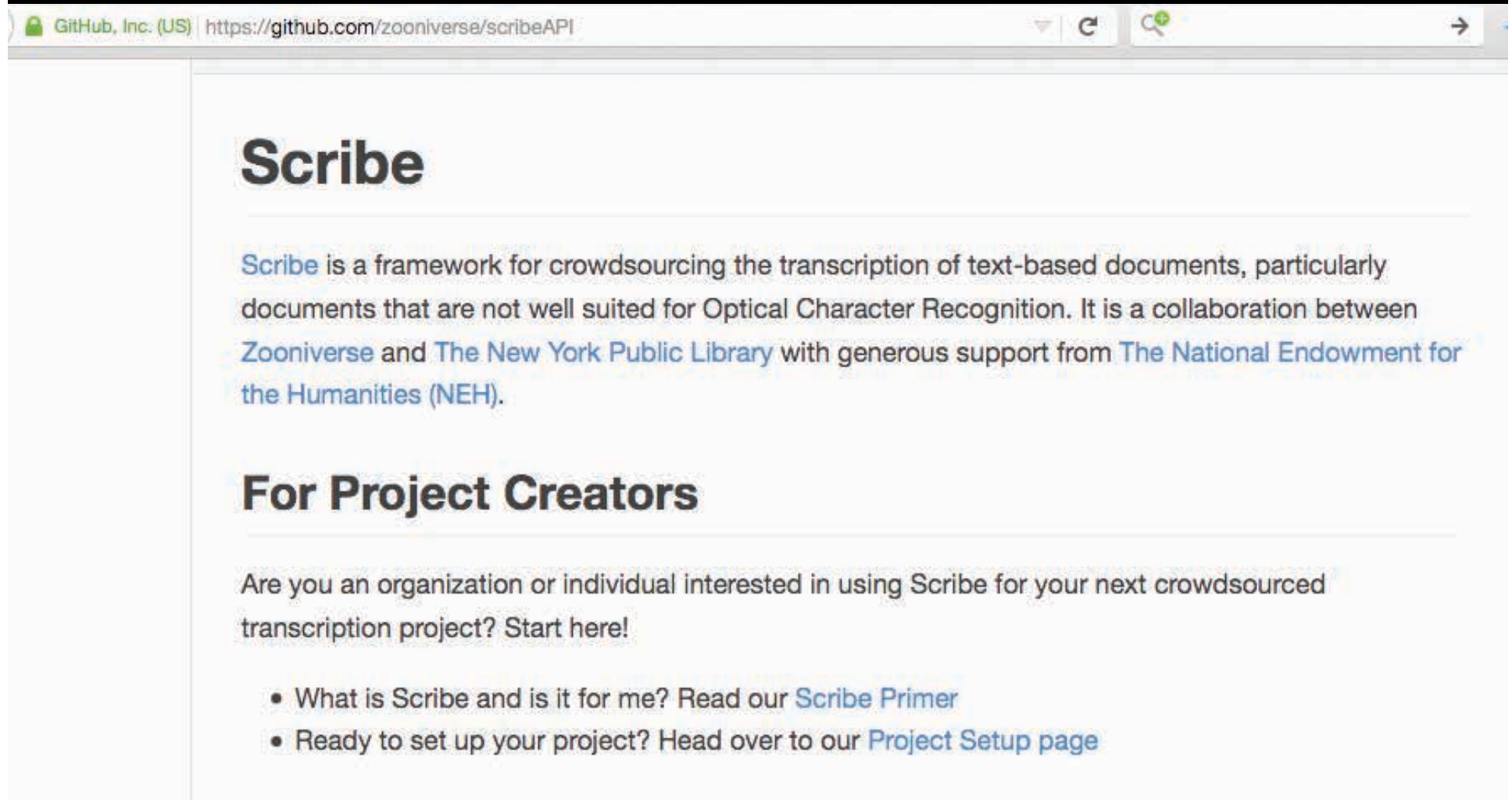
Future questions for our research team

- Effect of short-run events on long-term health
 - Russian flu epidemic (1889)
 - 1893 depression
 - Did the 1893 depression have different effects in different places
- Expanded sample linked to vital records
 - With history sheets, effect of wounds on later life health
- Link British-born troops to existing collections of complete British census records
- Better identification of Māori in sample with names of next of kin

Potential questions for other researchers

- Migration - where were people born, enlist, where were their parents?
 - Imperial and local connections
- Social networks. Who served with whom and how were they connected?
- Women in the war: 626 nurses + other auxiliary health care aides
- Māori / Pākehā inter-marriage – limited marital registration of Māori before 1913.
- Descriptions of the body: Tattoos, wounds
- Deviance and discipline from History sheets
- Structure of military careers

Potential for more demographic research



The image is a screenshot of a web browser displaying the GitHub repository page for the Scribe API. The browser's address bar shows the URL <https://github.com/zooniverse/scribeAPI>. The page title is "Scribe". The main content area describes Scribe as a framework for crowdsourcing the transcription of text-based documents, particularly those not well suited for Optical Character Recognition. It mentions a collaboration between Zooniverse and The New York Public Library, with support from The National Endowment for the Humanities (NEH). Below this, there is a section titled "For Project Creators" which asks if the user is an organization or individual interested in using Scribe for their next crowdsourced transcription project. It provides two bullet points: "What is Scribe and is it for me? Read our [Scribe Primer](#)" and "Ready to set up your project? Head over to our [Project Setup page](#)".

GitHub, Inc. (US) <https://github.com/zooniverse/scribeAPI>

Scribe

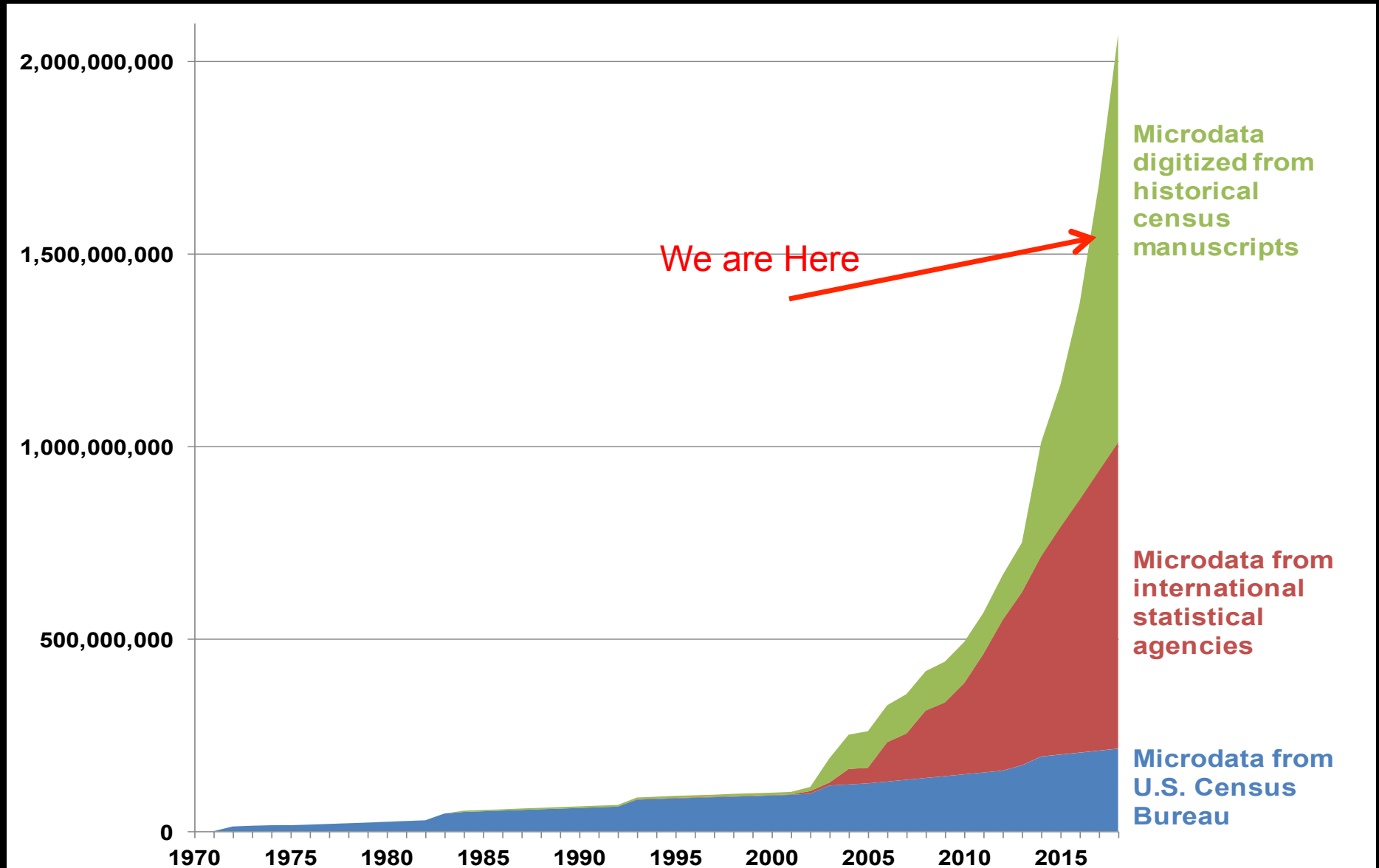
[Scribe](#) is a framework for crowdsourcing the transcription of text-based documents, particularly documents that are not well suited for Optical Character Recognition. It is a collaboration between [Zooniverse](#) and [The New York Public Library](#) with generous support from [The National Endowment for the Humanities \(NEH\)](#).

For Project Creators




Are you an organization or individual interested in using Scribe for your next crowdsourced transcription project? Start here!

- What is Scribe and is it for me? Read our [Scribe Primer](#)
- Ready to set up your project? Head over to our [Project Setup page](#)

Growth of public-use census microdata



PERSONS 14 YEARS OLD AND OVER—EMPLOYMENT STATUS																
OCCUPATION, INDUSTRY, AND CLASS OF WORKER																
For a person at work, designated as public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories: (a) Public transportation work, or with a job ("Job") in the following categories:																

 Rakefile	Demo	4 years ago
 config.ru	first commit	5 years ago
 license.txt	license.txt	4 years ago

README.md

Welcome to Scribe

Scribe is a framework for generating crowd sources transcriptions of image based documents. It provides a system for generating templates which combined with a magnification tool guide a user through the process of transcribing an asset (an image).

Getting Started

- We first need to install a mongodb server. This is then specified in config/mongodb.yml (see config/mongodb.hudson.yml for an example). Since databases are created lazily in MongoDB just specify the database name you want to use there.
- Site settings (config/site_settings.hudson.yml) contains the application name and other detail about the project. You should rename site_settings.hudson.yml to site_settings.yml
- To generate the templates for the project look at the lib/tasks/sample_weather_bootstrap.rake file. You need to specify each entity type you wish transcribed and its fields along with help text for the user for each.