

Crowdsourcing in libraries and museums: Challenges, opportunities and digital impacts

Dr. Mia Ridge, @mia_out
Digital Curator, British Library

Chinese Association of Museums
Taipei, Taiwan, August 2017

Thank you for the invitation to speak.

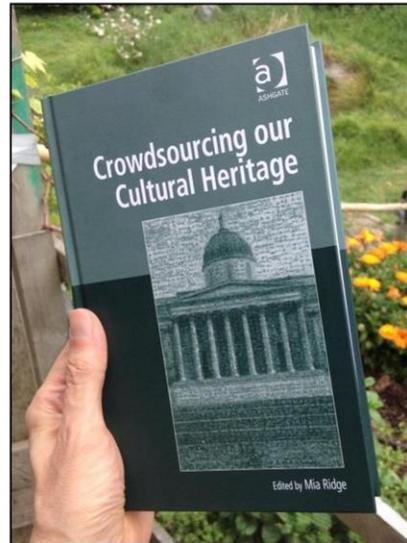
Abstract: Crowdsourcing in museums and libraries involves asking the public to help with tasks that contribute to a shared, significant goal or research interest related to cultural heritage collections or knowledge. This talk will introduce some key examples of successful crowdsourcing projects that have transcribed, categorised, linked and researched millions of cultural heritage and scientific records, and discuss some of the reasons for their success. Digital technologies have enabled exciting new forms of public participation in cultural heritage and the sciences, but they can also challenge museums, libraries and archives to manage the changes that these new opportunities bring. Audience expectations have changed as social media and digital technologies have encouraged greater organisational transparency, and more flexible digitisation and information management practices. How can museums and libraries work with the public to make collections more accessible while making room for people to explore and enjoy collections in new ways?

Overview

- Introduction: about me, definitions and key examples
- Opportunities for museums and libraries
- Challenges for museums and libraries
- Looking to the future
- Questions

About me

- Digital Curator, British Library
- Book 'Crowdsourcing our Cultural Heritage', 2014
- PhD 'Making Digital History: the impact of digitality on public participation and scholarly practices in historical research', 2015
- MSc in human-computer interaction: crowdsourcing games to enhance museum collection metadata, 2011



Why am I here talking about this? I've done major research projects on crowdsourcing in cultural heritage, and edited a book on the subject. My PhD in digital history studied crowdsourcing as a stepping stone to engagement with the practices and skills of history.

Dora's lost data

 "Hi, my name is Dora, and I'm a junior curator. It's my first day and I've made a big mistake – I accidentally deleted all the information we were going to add to our collections online. I need to re-label them, and quickly..."

Can you help? **Add words about the thing in the picture that would help someone find it on Google** – how it looks, what does, who might have used it – anything you can think of."

Scale model (1:10) of Swift gamma-ray burst satellite

Object from: Science Museum.
Date: c. 2000 Place: University of Leicester (Accession num: L2009-4034)



Add words ('tags') to describe this object

Tags

Tip: separate each tag with a comma, like this: tag, label, a phrase, name, names.

Not sure about this object? [Get a different object](#). (It won't affect your points.)

Tip: save this URL if you want more time to think or research: http://museumgam.es/dora/?obj_ID=225



I worked in social history and history of science museums for a long time, got interested in opportunity between public engagement and need to enhance collections. Made games for a research project in 2010, and as a result realised that a) games helped overcome fear of adding content in authoritative space, b) asking people to look at a series of scientific objects in order to describe them made them curious about the objects

Page Discussion
Read Edit View history Search

2nd Battalion Leinster Regiment, UK

This page was initially automatically created using data from series WO 95 (official war diaries) in the UK National Archives Discovery catalogue. It is not guaranteed to be complete or accurate. You can edit this page to make it more useful. Please help to improve the page by adding or correcting facts, and citing reliable sources.

Home

How you can help

Researching World War One records

About this project

Frequently asked questions and site help

Contact

Unit pages that need related diaries, letters

American units in World War I

Australian units in World War I

British units in World War I

British colonial units in World War I

Canadian units in World War I

French units in World War I

Indian units in World War I

Italian units in World War I

New Zealand units in World War I

Portuguese units in World War I

South African units in World War I

Diary pages that need links to units

personal diaries, letters, memoirs

official unit diaries and histories

Wiki stuff

Recent changes

Random page

Help

Tools

What links here

Special pages

Recent changes

Random page

Contents [hide]

- Personal narratives
 - John Staniforth
- Related media
- Official sources
 - Unit war diaries
 - Unit histories
 - Other official documents
- Other sources
- References

Personal narratives [edit]

John Staniforth [edit]

John Hamilton Maxwell Staniforth [en] enlisted as a private and rose to the rank of captain, serving in these units:

- 6th Battalion Connaught Rangers, UK
- 7th Battalion Leinster Regiment, UK
- 2nd Battalion Leinster Regiment, UK**

Imperial War Museum, Documents.14337 [en] is a collection of his private papers.

A selection of his letters has been published as Staniforth, John Hamilton Maxwell (2012). Grayson, Richard, ed. *At War with the 16th Irish Division 1914-1918: The Letters of J. H. M. Staniforth*. Barnsley: Pen & Sword. ISBN 978-1848846340.

Related media [edit]

Official sources [edit]

Unit war diaries [edit]

- The National Archives of the UK (TNA): WO 95/16122 [en] (01/08/1914-31/10/1915)

Unit war diaries [edit]

2nd Battalion Leinster Regiment, UK	
Years active:	unknown – unknown
Country:	United Kingdom
Service:	British Army
Branch:	Regular Army (British Army)
Type:	Infantry
Specific type:	Infantry battalion
Full size:	1,000
Sources for overview:	
Created:	unknown.
Sources for created:	
Name1:	2nd Battalion Prince of Wales's Leinster Regiment (Royal Canadians)
Name1 Start:	unknown
Name1 End:	unknown
Reason for change:	
Sources for name1:	
Alternate names:	
Sources for alternate names:	
Disbanded:	unknown.
Sources for disbanded:	
For more information on what infobox fields mean, see Documentation at military unit command structure and	

Also undertook more complex participatory project on social history of the First World War. Very niche project but has been successful in creating structured data about military units in WWI.

Digital Scholarship at the British Library



My team helps scholars use digitised, digital collections of the British Library in research and creative projects. Interested in application of computational methods in digital humanities, digital social sciences, etc, to our collections. We also provide advice on digitisation and run a training programme to teach foundational methods of digital scholarship.

Definitions

Crowdsourcing in cultural heritage

Asking the public to help with tasks that contribute to a shared, significant goal or research interest related to cultural heritage collections or knowledge.

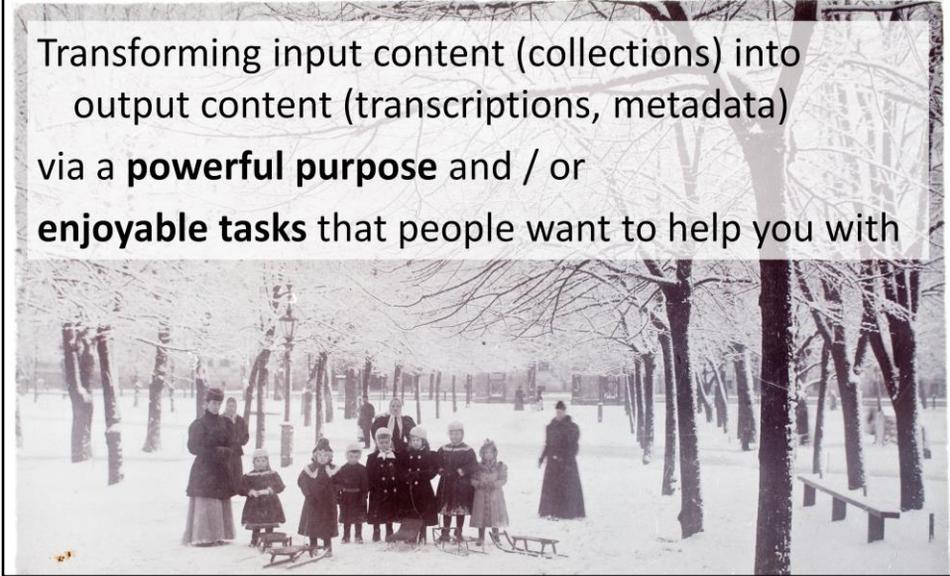
The activities and/or goals should be inherently rewarding.

My definition is partly descriptive, and partly proscriptive (what it should be, as well as what it is). Benefit should be wider than your institution e.g. improving catalogue data helps any user of the catalogue as well as the institution.

No financial rewards so has to be rewarding. Often task is quite enjoyable, and people are motivated by knowing their contribution helps make the world a better place.

Basically...

Transforming input content (collections) into output content (transcriptions, metadata) via a **powerful purpose** and / or **enjoyable tasks** that people want to help you with



Those terms 'powerful purpose' and 'enjoyable tasks' are key aspects of successful projects.

Input content - the image. Output content - terms to describe the image.

Heritage crowdsourcing as volunteering



'Online volunteering' is a good way of thinking about crowdsourcing in cultural heritage. Contributors are looking for a leisure activity - some just want casual activities they can pick up whenever suits them, others want an opportunity to develop deeper skills and interests. The opportunity to socialise with other people with similar interests can turn into a strong motivation for continuing for some volunteers.

If you've worked with in-person volunteer or community programmes, you already have a lot of the skills needed to run a good crowdsourcing project.

...but convenient



Digital tech offers serious advantages over in-person volunteering programmes. They are not tied to venue opening hours or location; not limited by conservation or handling issues once material is digitised. Allows you to reach thousands of people, or just a few interested specialists who might be located anywhere in the world. Convenience for volunteers means they can fit it in around their lives. A few minutes here and there adds up, means people can take up hobbies sooner (where previously they might have waited until retirement).

Satisfying 'microtasks'

The screenshot shows a web browser window with the URL 'A New York Public Library website'. The page title is 'What's On The Menu?'. The navigation menu includes 'Home', 'Menus', 'Data', 'Dishes', 'About', 'Blog', and 'iSearch'. The main content area displays a menu item 'Pineapple Water Ice' with a price of 25. A green arrow points to the dish name. Below the menu item, there is a form with the following fields:

- What does this say?
- Please type the text of the indicated dish EXACTLY as it appears. Don't worry about accents.
- Price (in Dollars): See more than one price for this dish?
- Trouble reading the text? Type your best guess, then check this box ->
-
- Delete this dish

At the bottom of the page, there is a footer with the following text:

PRIVACY POLICY • RULES AND REGULATIONS • POLICY ON PATRON-GENERATED WEB CONTENT • TERMS AND CONDITIONS
A PROJECT OF NYPL LABS, WITH GENEROUS SUPPORT FROM:
 **Humanities**
NYPL 'What's on the Menu?'

This is a screenshot of the Menus project. They have removed everything not essential to the task from the interface so you can focus on the task. They've anticipated common questions and provided answers on the screen.

Technology helps...
But ultimately
crowdsourcing is all
about people



James Murray, editor, OED, with contributor slips

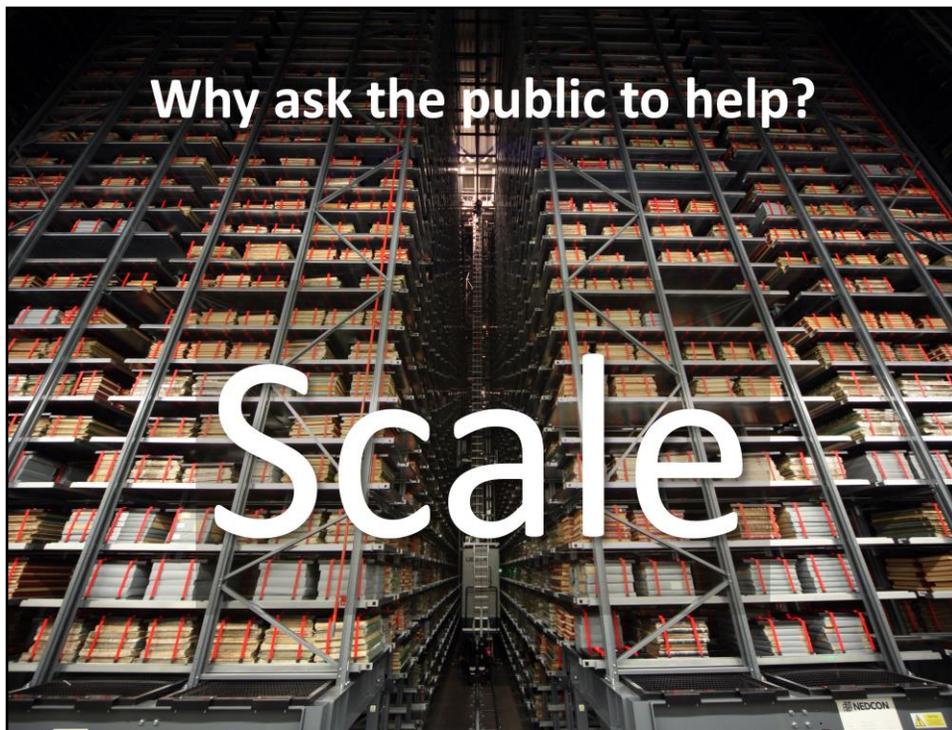
19thC examples of distributed networks of observations and research reported by post, telegraph... Technology has made it easier to reach interested people, to collate and verify work, but the basic model pre-dates computers. First example is the Oxford English Dictionary (OED). In 1879 the editor asked the public to help with finding sources and the earliest examples of use for words for which they were lacking information. Contributions from the public would be used in the dictionary alongside those from recognised and invited experts.

This shows two things - even then, a call for public assistance could be embedded in scholarly practice, where the same intrinsic motivations of leisure, social networks and community, learning, an interest in the subject and a chance to practice skills were important. You can still participate in OED appeals today, so it must work.

It also shows that crowdsourcing as we know it has been transformed by technology, but not created by it. The ability of digital technology to provide almost instant data gathering and feedback, automatic validation and the ability to reach both broad and niche groups through loose networks have all been particularly important. For collecting institutions, technology has also helped manage the sheer physical issue of providing access to collections without space or conservation limitations.

See also: <http://blog.oxforddictionaries.com/2013/02/james-murray/>
<http://public.oed.com/history-of-the-oed/archived-documents/april-1879-appeal/april-1879-appeal/>

Why crowdsource cultural heritage work?



Digitisation backlog: collections are big, resources are small. Manually enhancing collections records is expensive and time-consuming. Very few orgs have the resources for digitisation of images, let alone full cataloguing or transcription.

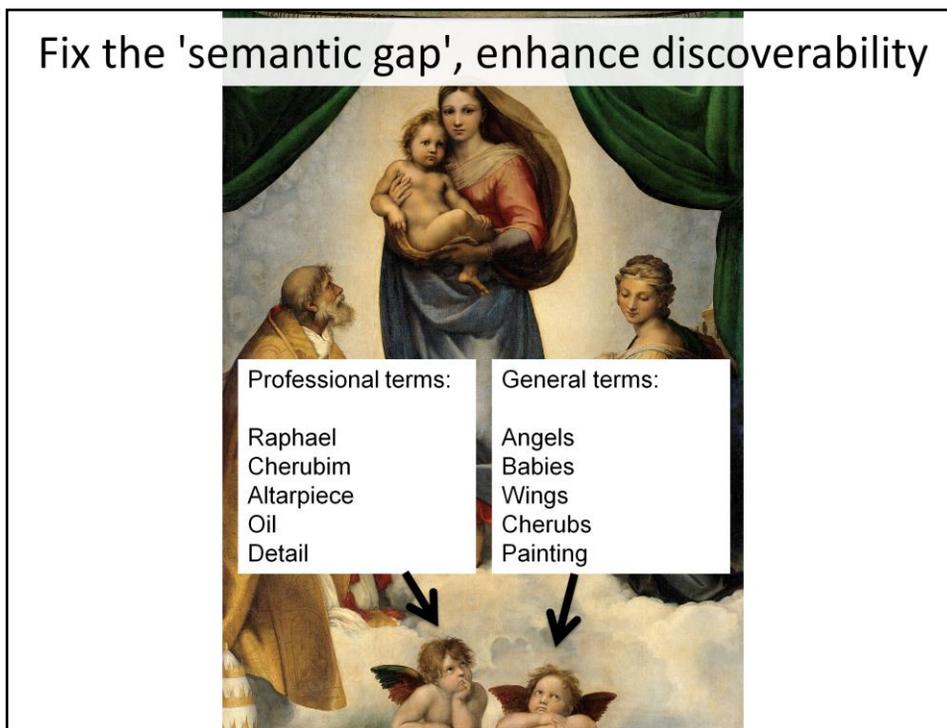
The British Library holds 180 - 200 million items:

Over 14 million books; 8 million stamps; 310,000 manuscript volumes; 4 million maps; 60 million patents; 260,000 journal titles; sound files; pamphlets, magazines, sheet music, newspapers; television and radio recordings; websites

Over 3 million new items are added every year

Only 1-4% of collections are digitised/digital

Image: The storage void of the new British Library National Newspaper Building at Boston Spa in West Yorkshire. Photo © Kippa Matthews



Discoverability is the ability of something to be found. Detailed metadata, descriptions and images help make items discoverable.

However, when metadata or information records are created by professional cataloguers, the content is often designed for internal or specialist users and doesn't use the everyday language our audiences might use to find material. This example shows the difference between the language a cataloguer and a member of the public might use to describe the cherubs in this painting.

Image: Raphael - Madonna Sixtina (Gemäldegalerie Alter Meister, Dresden, 1513-14

Access external specialist expertise



There's a lot of specialist expertise outside the museum. There's an online community for almost every topic or type of item under the sun, and some of those people might want to share their knowledge with a museum or library. People can share information about their professional or social experiences of a museum object, or share the results of their personal research.

Create meaningful experiences with collections



Well-designed projects can help people discover new interests, communities, or just encourage them to have a brief moment of deeper engagement with cultural heritage. Will talk more about that later.



Might be learning by following their own interests using heritage material, through discussion with others and with museum staff, or just through spending lots of time developing familiarity with material.

Key examples (and why they work)

Types of projects

'Type what you see'

'Describe what you see'

'Share what you know'

'Validate other inputs'

About what the task involves. The uses of the data, and accuracy and validation requirements are different for each type.

'Type what you see'

Sometimes literally 'type the text that's in front of you', other times it's an action based on what you can see on the screen.

reCAPTCHA

The image shows a screenshot of the reCAPTCHA website. On the left, there is a large red arrow pointing downwards and to the right. The reCAPTCHA logo is at the top left. A navigation menu includes: WHAT IS reCAPTCHA, GET reCAPTCHA, PROTECT YOUR EMAIL, MY ACCOUNT, and RESOURCES: DOCS & PLUGINS. The main heading reads: reCAPTCHA IS A FREE ANTI-BOT SERVICE THAT HELPS DIGITIZE BOOKS. Below this is a snippet of text from a scanned document: 'steamboat train, from New York this morning' with a red box around the word 'morning'. A red arrow points from the word 'morning' in the text to a reCAPTCHA widget. The widget shows the text 'morning' and 'upon' with a red box around 'morning'. Below the widget is a red button that says 'USE reCAPTCHA ON YOUR SITE'. To the right of the button are three icons with text: a shield for 'STRONG SECURITY', a person with a plus sign for 'ACCESSIBLE TO BLIND USERS', and a bar chart for '30+ MILLION SERVED DAILY'. At the bottom, there is a small text block: 'See how accurate reCAPTCHA is at digitizing content!' and a footer with links: 'Blog | About Us | Contact | FAQs | Terms' and '© 2013 Google. All rights reserved.'

You may have seen something like this when commenting or registering for a site online. If you've typed in the words shown, then you've helped correct text (or transcribe audio). A 'type what you see' example.

Trove: OCR correction

The screenshot displays the Trove website interface for a newspaper article. At the top, it says "NATIONAL LIBRARY OF AUSTRALIA" and "Trove". The article is from "The Canberra Times (ACT) : 1926 - 1995" dated "Fri 31 Oct 1969", page 15, titled "TAIWAN TOURIST SURVEY".

The article text is partially visible, with some words highlighted in green, indicating OCR corrections. The main heading is "OF Fascination". Below the heading is a photograph of Chiang Kai-shek, the President of the Republic of China, with the caption "Chiang Kai-shek, the President of the Republic of China.".

The left sidebar shows a list of items related to the article, including "TAIWAN", "TOURIST SURVEY", "A concentration of fascination", and "THE Island of Tai".

The top of the article text reads: "THE island of Taiwan is not very big. It is approximately 240 miles long, 85 miles wide, and 13,808 square miles in area." The text continues with a description of the island's geography and tourism.

National Library of Australia's Trove interface for newspapers lets people correct errors they find when looking through historical newspapers. The task is so satisfying that some people spend hours a week on it. They have minimised barriers to participation - if you see an error, you can 'fix this text'.

Designed to let people get on with correcting errors they'd come across when doing their work, but satisfying enough task that people do it for fun. Minimised barriers to participation. 238 millions lines of text correct

Rewards reinforce motivation

Date modified	Old Lines	New Lines
a minute ago miah	when they first sighted it in 1590, means sie Beauti	when they first sighted it in 1590, means Isle Beauti
a minute ago miah	of fascination THE island of Tai wan is not very big. It is approximately 240 mile long, 85 miles wide, and 13,808 square miles in area. Yet it is a fascinating place for the tourist and it is a most point whether the small size of the island, facilitating as it does travel from place to place. The natural scenery is both superb and varied. Mainly tropical, it has lowering mountains, deep valleys and lovely lakes. The original name, Ilha Formosa, bestowed on it by Portuguese sailors when they first sighted it in 1590, means Isle Beauti and the visitor will appreciate the aptness of it. But if nature has been bountiful to Taiwan, man has added excitement	of fascination THE island of Tai wan is not very big. It is approximately 240 miles long, 85 miles wide, and 13,808 square miles in area. Yet it is a fascinating place for the tourist and it is a most point whether the small size of the island, facilitating as it does travel from place to place. The natural scenery is both superb and varied. Mainly tropical, it has lowering mountains, deep valleys and lovely lakes. The original name, Ilha Formosa, bestowed on it by Portuguese sailors when they first sighted it in 1590, means Isle Beauti and the visitor will appreciate the aptness of it. But if nature has been bountiful to Taiwan, man has added excitement

Your contribution makes a difference immediately... Effective design that makes correcting text a satisfying interaction. The user experience is further enhanced by the immediate appearance of the corrected text on the page (alongside the editing history). This shows participants the value of their contribution by making their corrections immediately available for the benefit of other users, reassures them that their work can be checked.

ary website [Explore otherst](#)

NYPL Labs  **What's on the menu?** Est. 2011

[Menus](#) | [Dishes](#) | [Data](#) | [Blog](#) | [About](#) | [Help](#)

Help The New York Public Library improve a unique collection!

We're transcribing our historical restaurant menus, dish by dish, so that they can be searched by what people were eating back in the day. It's a big job so we need your help! [Learn more.](#)

Connect: menus@nypl.org | [Twitter](#) | [Facebook](#)

We need you!

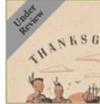
Help review

It's easy! No registration required!

So far: 1,331,936 dishes transcribed from 17,545 menus

Giving Thanks

A look at some of our favorite Thanksgiving menus

						
Hotel Normandie 1889 62 dishes	Briggs House 1899 20 dishes	Gramercy Park Hotel 1949	Occidental Hotel 1891 40 dishes	Hotel Rodman 1905 42 dishes	Hotel St. James 1899 37 dishes	Map our Menus! Need a break from transcribing? Help us locate menus with our new Geotagger tool.

New York Public Library's What's on the menu project was a very successful example. Transcribed dishes and prices from 17,000 menus.

Works because it has a really focused design. It's also regularly updated (e.g. topical menus - spring, Superbowl), showing that the organisation is paying attention and values contributions. Points people to content made discoverable through the project.

NYPL Labs **What's on the menu?** Est. 2011

Search keyword(s)

Menus | Dishes | Data | Blog | About | Help

Help The New York Public Library improve a unique collection!

We're transcribing our historical restaurant menus, dish by dish, so that they can be searched by what people were eating back in the day. It's a big job so we need your help! [Learn more.](#)

Connect: menus@nypl.org | [Twitter](#) | [Facebook](#)

Frutti di Mare!

Latest catch of menus...

Fisherman's Grotto 6 dishes	The Famous Anchor Sea Food House 1962 358 dishes	Rogano Restaurant & Sea Food Bar 1964 21 dishes	Anthony's Fish Grotto 184 dishes	American Restaurant 1987 99 dishes	Anchor Inn Seafood Restaurant 1968 175 dishes

Map our Menus!

Need a break from transcribing? Help us locate menus with our new Geotagger tool.

Help review

Transcribed menus that need a second pair of eyes. Help fix misspellings, fill in missing data...

Pennsylvania Railroad Dining Cars 1900 39 dishes	Dea The Beachcomber 1965 169 dishes	The Orient Room 1964 982 dishes	American Cotton Manufacturers' 1907 15 dishes	Zum Ruten Baeren 1939 27 dishes	Waldorf Astoria 1912 18 dishes

Explore

Browse the collection. 17,545 menus digitized and counting...

Today's specials

An assortment of dishes from the collection

- Fresh Pear Salad
- Farina Jelly With Cream
- Seaf Gurken
- Cold Tongue
- Edam Cheese
- Stewed Fruit, Pears
- Peach Melba
- Anisette Cordial
- Lamb, Mint Jelly

WEB FONTS BY fonts.com

It taps into altruistic and subject specialist motivations. Gives a clear sense of what to do next - literally has hands pointing to the button to press.

Like Trove, they have removed text or processes that are barriers to participation. In design terms, they've removed the 'friction' from getting started. They've also anticipated worries like 'do I need to register?' and answered them right there on the page. Overall, it shows that empathy for your participants is an important design asset.

FamilySearch Indexing

FOR INFORMATIONAL PURPOSES ONLY

Batch ▾ Data Entry ▾ Help ▾
QA Brad -- Image Set Holds One [MMML-GFZ]
M Ridge Indexing

Saved: 8:15:36 pm

Field Help [\[\[en\]\]](#) [Principal Given \(MH\)](#) [More](#)
Step 2: Records ▾

Image 1 of 5

Record 1 of 1

[[en]] Principal Given (DN) *

[[en]] Principal Family (DN) *

[[en]] Principal Birth Date (DN) *

[[en]] Father Given (DN)

[[en]] Mother Given (DN)

New Record

PLACE OF DEATH

County *Sanders* State *Mont* Registered No. *72106*

Village *Deer Lodge*

FULL NAME *Ida Blanche Clark*

SEX	COLOR OR RACE	Single, Married, Widowed, or Divorced	DATE OF DEATH
<i>female</i>	<i>white</i>	<i>Married</i>	<i>11 5 1918</i>

HUSBAND of *Paul Clark*
WIFE of *Paul Clark*

DATE OF BIRTH *1 27 1899*

AGE *19* Years *10* Months

OCCUPATION OF DECEASED *Housewife*

BIRTHPLACE *Ida*

NAME OF FATHER *Albin Stewart*

BIRTHPLACE OF FATHER *Ida*

MAIDEN NAME OF MOTHER *Mary Butler*

BIRTHPLACE OF MOTHER *Ida*

I filed this certificate for record in this office on *4-10-19*

County Clerk and Recorder of Sanders County, State of Montana

I HEREBY CERTIFY, That I attended deceased from *Oct 27* *1918* to *Nov 5* *1918* that I last saw h. *at* alive on *Nov 5* *1918* and that death occurred, on the date stated above, at *7 P. M.*

The CAUSE OF DEATH was as follows:
Chronic Bronchitis

CONTRIBUTORY *Influenza*

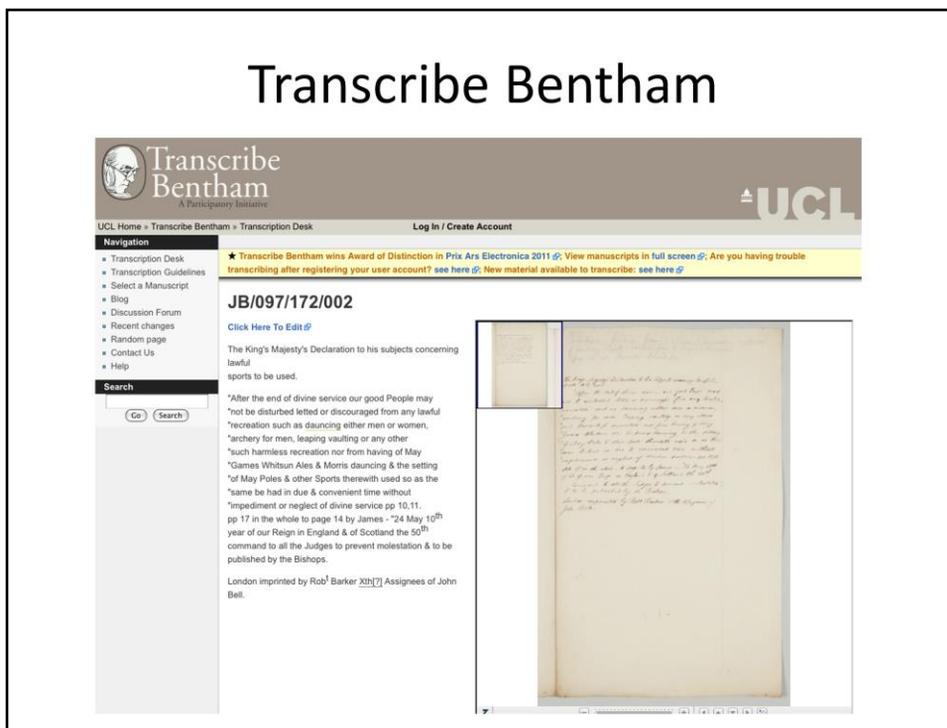
(Signed) *Ida Blanche Clark* M. D.

Place of Burial, Cremation, or Removal _____ Date of Burial _____

This handwriting transcription task is harder and takes more time, but people are highly motivated to contribute to family history (and for some people, it's a contribution to their religious mission).

How easy transcription tasks are depends on the handwriting, but that's part of the attraction

Transcribe Bentham



The screenshot displays the Transcribe Bentham website interface. At the top, the site logo features a portrait of James Bentham and the text "Transcribe Bentham A Participatory Initiative". The UCL logo is in the top right corner. Below the header, a navigation menu lists options like "Transcription Desk", "Transcription Guidelines", "Select a Manuscript", "Blog", "Discussion Forum", "Recent changes", "Random page", "Contact Us", and "Help". A search bar is also present. The main content area shows a transcription task for the document "JB/097/172/002". The task title is "The King's Majesty's Declaration to his subjects concerning lawful sports to be used." Below the title, there is a "Click Here To Edit" link. The document text is displayed in a large font, with a small inset image showing a zoomed-in view of the original manuscript page. The text includes a quote: "After the end of divine service our good People may not be disturbed letted or discouraged from any lawful recreation such as dauncing either men or women, archery for men, leaping vaulting or any other such harmless recreation nor from having of May Games Whitsun Ales & Morris dauncing & the setting of May Poles & other Sports therewith used so as the same be had in due & convenient time without impediment or neglect of divine service pp 10,11. pp 17 in the whole to page 14 by James - 24 May 10th year of our Reign in England & of Scotland the 50th command to all the Judges to prevent molestation & to be published by the Bishops. London imprinted by Rob^t Barker Xih^t Assignees of John Bell.

Complex task - marking up transcriptions in XML - on difficult source material. Has a small number of very productive super-taggers... manual validation creates backlog and delay in approving content reduces feelings of reward. Post to the blog about progress help make up for it. Media coverage helped - each round drew in a few contributors.

Smithsonian HOME PROJECTS ABOUT TIPS NEWS

SMITHSONIAN DIGITAL VOLUNTEERS: TRANSCRIPTION CENTER SIGNUP LOGIN

FEATURED
U.S. NATIONAL HERBARIUM
VIEW PROJECTS

JOIN US!
LEARN HOW TO TRANSCRIBE
Become a Smithsonian Digital Volunteer and help us make historical documents and biodiversity data more accessible. Join 5,592 volunteers and contribute to field notes, diaries, ledgers, logbooks, currency proof sheets, photo albums, manuscripts, biodiversity specimens labels and more - [Get started Now!](#) Follow us on Twitter and learn more about projects: [@TranscribeSI](#) Partner on projects and ask your [#volunteers](#) for best tips and tricks.

BROWSE PROJECTS
Select a category below to begin browsing projects.
Select a Category

BE A PART OF IT
New York Certified Proofs (Box 1)
638 CONTRIBUTIONS
Which cities will you find in the first box of certified proofs from New York, part of the National Numismatics Collection?
[New York Certified Proofs \(Box 2\)](#)

LATEST UPDATES
J.Proctor transcribed a page from The Legend of Sgenhadishon, 1920-1928
J.Proctor transcribed a page from The Legend of Sgenhadishon, 1920-1928
andk marked for review a page from Dove Diaries, 1942-3
andk transcribed a page from Dove Diaries, 1942-3
margrace transcribed a page from Arthur W. Stelfox - Vol. 2, daily lists
andk reviewed a page from Dove Diaries, 1942-3

feedback

Relatively recent project, was able to learn from previous projects.

Revel in niche topics and challenges



21% COMPLETE

41 Total Pages
12 Contributing Members

START CONTRIBUTING TODAY.

BAILEY - FIELD NOTES, MONTANA, WASHINGTON, AND OREGON, 1909

Do you have an eye for detail? Bureau of Biological Survey special field agent Vernon Bailey (1864-1942) took careful notes of the plants and animal specimens he observed as he traveled. In 1909, while enroute to Gold Beach, Oregon from Montana, he kept his notes in this field book. Please help us transcribe these notes to make them more accessible for future researchers and scholars. You might want to look at other transcribed [Bailey field notes](#) to get a sense for his handwriting style.



74% COMPLETE

126 Total Pages
32 Contributing Members

START CONTRIBUTING TODAY.

BOHUMIL SHIMEK - DIARY, EUROPEAN TRIP, 1914 (1 OF 2)

** If you can read Czech, we need you! Help us and other digital volunteers to finish transcribing Shimek's first volume of notes from his trip to Europe on the eve of World War I. ** Practicing a new language on your first trip to Europe? Surveyor-turned-naturalist Bohumil Shimek was working on Czech when he traveled from Iowa to Prague as an exchange professor of botany in 1914. Just before open hostilities broke out in what became known as the Great War or World War I. A son of Czech immigrants, Shimek came to be well known for his contributions to the field of botany in North America - over 205,000 plant specimens. Take note, he writes in English and Czech! [English speakers may want to start here](#), a few pages in. You also might enjoy this [blog post](#) about his trip.

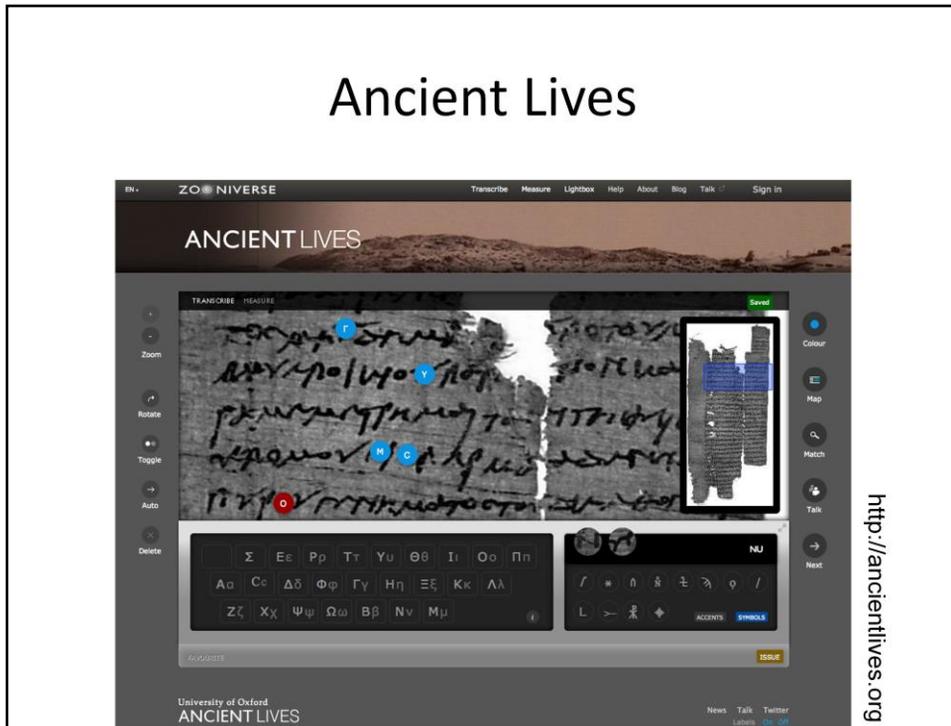


22% COMPLETE

JOURNAL KEPT BY V. BAILEY ON FIELD TRIP TO MICHIGAN AND

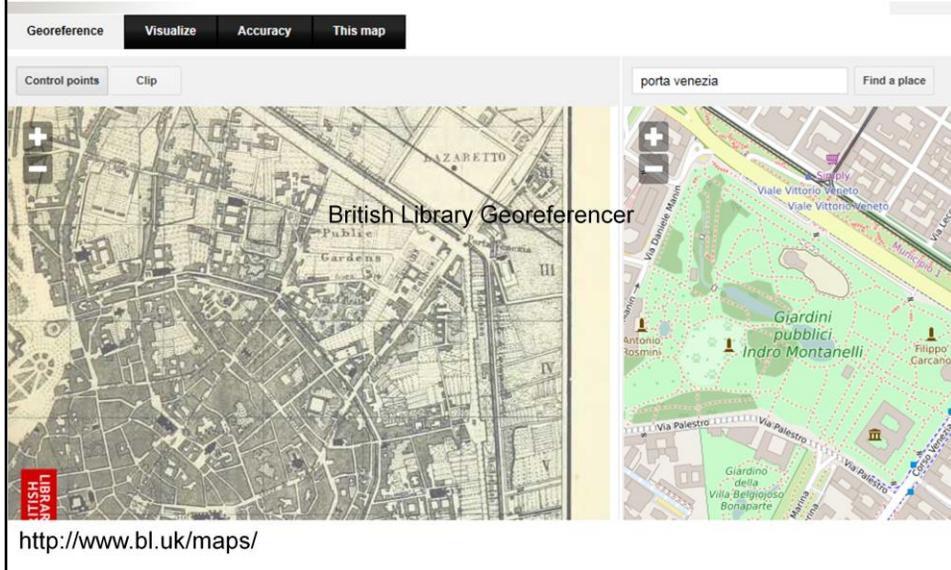
The design breaks the overall task of transcribing Smithsonian's collections into 'projects' the size of a notebook. This means each 'project' feels easier to complete, and gives them lots of successes to celebrate. They are able to provide a description for each notebook that includes specific places, languages, people that might attract your interest.

Ancient Lives



Possibly a great example for students - transcribe Ancient Greek letters from papyri, even if you can't read the Greek alphabet. When trying this out I found myself slowly getting better at spotting different Greek letters because the interface breaks it down into a 'pattern matching' task.

British Library Georeferencer



Moving on to tasks that involve visual matching... In this British Library project, people match historical and modern maps and add markers to each to align historical maps with modern coordinate systems. This means that old maps can be displayed spatially. It's an enjoyable task if you like maps and solving puzzles, and a great way to interact with the collections.

British Library Sound Maps

Sound map - UK Soundmap

The UK Soundmap, the first nationwide sound map, invited people to record the sounds of their environment, be it at home, work or play. Over 2,000 recordings were uploaded by some 350 contributors during the period July 2010 to July 2011. [Further information](#).



Your accents <http://sounds.bl.uk/Sound-Maps/Your-Accents>
<http://sounds.bl.uk/Sound-Maps/UK-Soundmap>

This is an example of a 'collecting' project, and also an example of an audiovisual project. People around the UK could record local sounds and upload them. They had to be checked in case copyrighted material like music was playing in the background, but some people were delighted when their contribution was accepted.

Micropasts photo masking

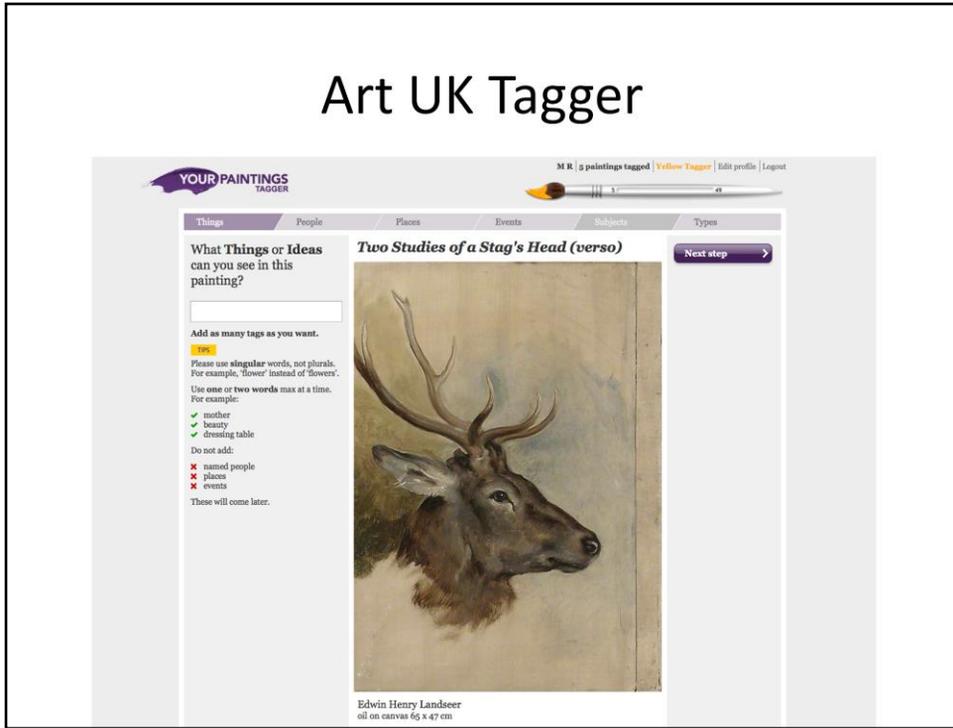


A lot of the tasks already mentioned are versions of traditional cataloguing tasks. This Micropasts 'photo masking' task is entirely new. They have provided a series of photos of objects taken from different angles. You draw around them so the computer can create a 3D model of the object.

'Describe what you see' /
'Share what you know'

The line between adding tags that describe something and adding information drawn from knowledge beyond what's on the screen is blurry.

Art UK Tagger



Tagging – typing in words that describe what you see – can be useful and give people an excuse to spend time with collections. Immersion in images can be its own reward.

Flickr and Flickr Commons



Is anyone using Flickr Commons for sharing photos?

SSO... + Follow

5,481 views 15 faves 6 comments Taken on October 24, 2005
No known copyright restrictions

Show EXIF

This photo is in 1 group

"History & Antiquities"
(Post 5 - Award 3)
282,830 items

This photo is in 1 album

Nantucket Brides
20 items

Tags

Nantucket Weddings
Brides woman portrait
lace
Kozorek Studio (New Bedford)
Almira (Flossie) West Williams
Almira Floretta West
Mrs. Charles H. Williams
wedding Wampanoag
Chappaquiddick
commons:event=common...

WoW photos back then are just amazing to look at.

Jun Acullador 6y
Hi, I'm an admin for a group called "History & Antiquities (Unlimited Posting - Award 3)", and we'd love to have this added to the group!

Pixel Wrangler 6y
Almira West Williams is the daughter of John E. West and sister to Gertrude West Ames.

wealthywamp 6y
This is my relative. She also went by the name of Flossie. She was Chappaquiddick Wampanoag. Her Grandmother was Elizabeth Brown West the daughter of Lucy Wamp Brown and Abram Brown who lived on the Indian reservation lands of the island of Chappaquiddick. Her great-grandmother was Wealthy Wamp. Almira (Flossie) lived in Newbedford.

People can add tags or comments to images posted to Flickr. Words like 'brides' help make the photos more discoverable, while two commenters have provided possible identification for the woman in the photo.

Tags - describe what you see; Comment - share what you know



10 Most Wanted
a website for the Museum of Design in Plastics

[Login](#)
[About](#)
[Blogs](#)

[How to Begin](#) | [Points & Prizes](#) | [Hall of Fame](#) | [Research Tips](#) | [Solved Cases](#) | [Cold Cases](#)

Home

Ice bucket and tongs

This is a striking object about which we know little other than it was made in Australia. We are hoping the down under members of the Plastics Historical Society might help us to learn more about it.



Case number - AIBDC : 0_6445

A turquoise, white and silver coloured ice bucket and tongs set. The ice bucket has a dribbled silver pattern and an inner liner with drainage holes in the bottom.

[View more images on the MoDIP site](#)

Designer	Unknown - Wanted
Manufacturer	Unknown - Wanted
Country	Australia
Date	1900 - 1990 - Wanted
Dimensions	height 150 mm, diameter 140 mm
Materials	plastic, unidentified - Wanted
Method	unknown - Wanted
Colours	blue, silver, white
Inscription	"Made In Australia"

Images on this site are for non-commercial, educational use only. MoDIP has done its utmost to obtain clearance from all IPR holders before adding images to this catalogue, if you believe that any image has been used without permission please contact us on modip@aub.ac.uk.

[Share](#) | [Tweet](#) | href="http://10most.org.uk" | [Join the hunt at: 10mostinvolved](#)

[f](#) | [t](#) | [g+](#) | [s](#)

Facebook Feed

Kuhle Beer AIBDC0_6445 Marble effect ice bucket with matching blue tongs 22/08/13

Phil Blume That is cool, I'm guessing 1950s, maybe 60s? We know it was made in Australia, so we're definitely going to need help with this one. <http://10most.org.uk/artefact/ice-bucket-and-tongs> 11/10/13

Phil Blume To remind users that interesting material posted here may be copied 27/11/13

Case notes

Australian connection
Submitted by **Phil Blume** on Fri, 11/10/2013 - 12:59.

Has anyone got connections to a friend in Australia who may be able to help here?

An experimental project. Museum of Design in Plastics tried to use crowdsourcing to find information about designers, methods, manufacturers for specific objects. Much more specialist knowledge or research skills required to contribute.

When asking people to contribute expertise or knowledge, can take longer for the right person to find the right question/object.

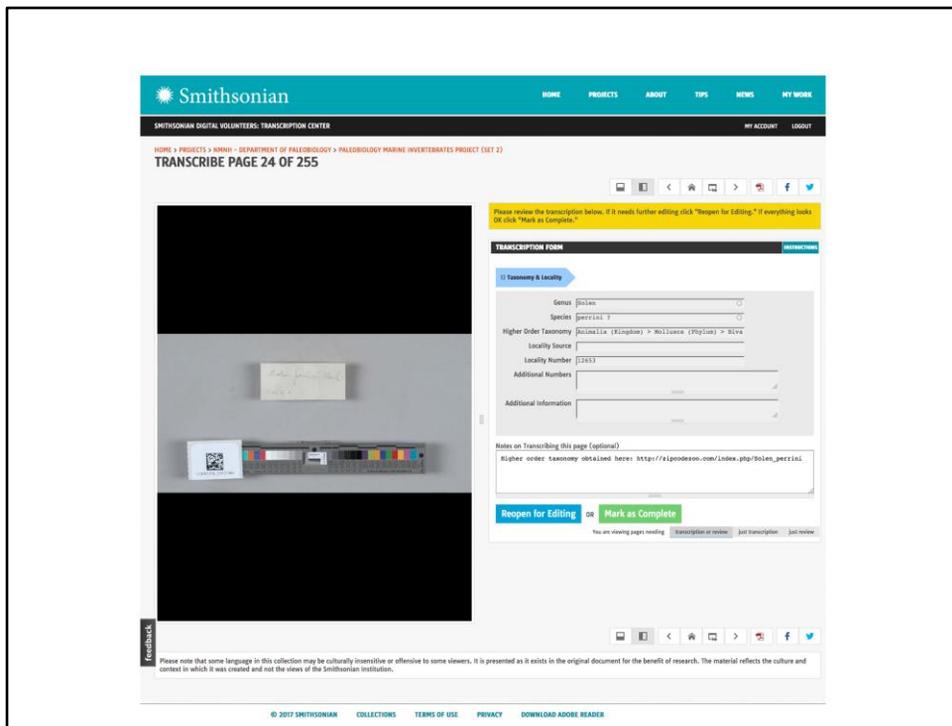
'Validate other inputs'

About what the task involves. The uses of the data, and accuracy and validation requirements are different for each type.

The screenshot displays the 'PCF Image Recognition Check' interface. At the top left, logos for 'BBC', 'pcf', and 'PUBLIC CATALOGUE FOUNDATION' are visible. The title 'PCF Image Recognition Check' is centered at the top. Below the title, it indicates 'Page Number 3 of 20' and 'Showing images 101 to 150'. The subject is 'sideburn'. Navigation buttons include 'Go to page', 'Previous Page', 'Make Page Green', 'Make Page Red', 'Next Page', and 'Save Page'. The main area shows a grid of 10 portrait paintings, each with a colored border (green, yellow, or red) and a caption below it: '101. Link to Your Paintings', '102. Link to Your Paintings', '103. Link to Your Paintings', '104. Link to Your Paintings', '105. Link to Your...', '106. Link to Your Paintings', '107. Link to Your Paintings', '108. Link to Your Paintings', '109. Link to Your Paintings', and '110. Link to Your...'. The interface is designed for quick review and classification of images.

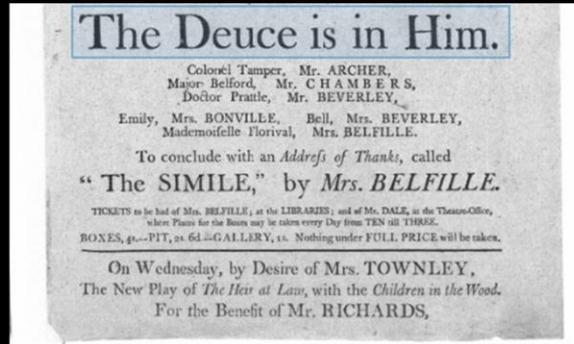
In this interface, people check suggested categories applied by software. You ignore any images that correctly match the category 'sideburn', click once on images that don't include sideburns, and twice on images that are uncertain. The design makes it easy to review lots of images very quickly.

This is an example of a workflow that combines people and software contributions - we will see a lot more of this in future.



The Smithsonian project asks registered volunteers to review other people's transcriptions for correctness. This reduces the amount of work that the organisation has to do to finalise records.

Designing task ecosystems



Validation tasks can be part of an 'ecosystem' of tasks that collectively contribute to getting all the needed data collected, checked and published.

**Opportunities for museums and
libraries**

Defining success: productivity

Text correction hall of fame

Automatically extracting text from scans of old newspapers and magazines is extremely challenging. Although the best available Optical Character Recognition (OCR) software has been used, the condition of the images it has to process combined with the frequently small fonts used means that many errors of interpretation are made.

Thankfully, many people have stepped in to correct the text and in so doing have made a wonderful contribution to this resource, and helped to improve search results. The following table lists by month the people who've corrected the most lines of text.

Total number of lines corrected for all time : 238,520,546

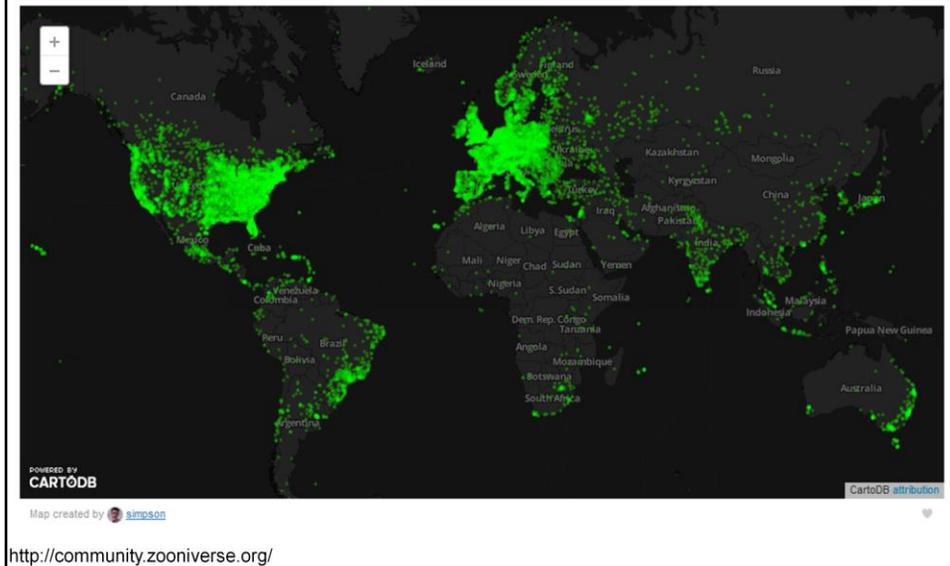
Leaderboard

All time Newspaper articles View My rank: 4,390th / total 48,345

Rank	Username	Lines corrected
1	JohnWarren	4,573,230
2	NeilHamilton	3,091,483
3	noelwoodhouse	3,049,100
4	John.F.Hall	2,276,454
5	anormanley	2,262,888
...
4,387	sheerolls	3,484
4,388	HelenCooper	3,482
4,389	stace	3,482
4,390	mlah	3,481
4,391	c.strom	3,479
4,392	deb31utante	3,479
4,393	Peter.Mairose	3,479
...
48,345	snoopy500	0

You can define success by the number of tasks completed. People will work towards metrics you provide, so be careful! If you focus on the number of tasks, people may rush and neglect the quality of a task, or not have the same level of absorption in a collection.

Defining success: reach



You can also define success by the number or type of people engaged in the tasks. This is often important for organisations whose mission is to reach the public, whether that's to give them an experience of contemporary science or access to their history through specific collections. The Zooniverse projects have reached well over a million people worldwide - this map. Some of scientific projects might also look at the impact of publications on social media and in journals that result from their projects. Museums might look at the number of researchers who find their digitised collections.

Defining success: engagement



Harry Corbyn Levinge

Mr **Harry Corbyn Levinge** (1831-11/1896) of Knock Drin Castle, Mullingar, J.P., D.L. Spent many years in the Indian Civil Service (Bengal), where he collected ferns, particularly in Sikkim and Kashmir. Unfortunately much of his collection was destroyed in a fire. On his return, he interested himself in the Irish flora and contributed a number of important papers on this topic to the Irish Naturalist and the J of Botany (1892). Published "The plants of Westmeath" in 1894. The fern *Adiantum levinge* was dedicated to him by J.G. Baker. His herbarium is now housed in the National Museum in Dublin.

WESTMEATH PLANTS.
By E. F. AND W. R. LASTON.

The plants here enumerated are most of them records for District VII. of the *Cybele Hibernica* (these have an asterisk prefix), and were observed by us while enjoying the hospitality of Mr. H. C. Levinge at Knock Drin Castle last July. Some of them were detected by the Rev. E. S. Marshall, our fellow-guest for part of the time, and some were pointed out to us by Mr. Levinge himself, whose investigations have done so much to throw light on the Flora of Westmeath. A rainy season had set in, and hindered work during part of our visit; but, thanks to the long preceding drought, many wet localities were unusually accessible, and all the longes had sunk below their normal level.



1364 *Ep. herb. H. C. Levinge*

Adiantum intacta Bull.

near Ballyvaughan. Co. Clare.

Ep. herb. H. C. Levinge May 18. 1892

Further info : see

<http://herbariaunited.org>

Finally, you can look at the number of people who are deeply engaged - people whose feelings or knowledge about the material or the underlying disciplines change to the extent that they change some aspect of their behaviour. In this example, people came for the herbarium specimens and got caught up in biographical interest of the original specimen collectors. There's something about handwritten text that seems particularly likely to get people thinking about historic lives.

http://herbariaunited.org/wiki/Harry_Corbyn_Levinge or
http://herbariaunited.org/wiki/Augustin_Ley

Crowdsourcing as 'stepping stones'

The screenshot shows the 'herbaria@home' website interface. On the left is a navigation menu with options like 'Herbaria@home', 'Message board', 'Links', 'Search', 'Specimens', 'Collectors', 'Taxon names', and 'VOCB match'. The main content area features a profile for 'Harry Corbyn Levinge' with a bio: 'Mr Harry Corbyn Levinge (1831-11/1896) of Knock Drin Castle, Mullingar, J.P. D.L. Spent many years in the Indian Civil Service (Bengal), where he collected ferns, particularly in Sikkim and Kashmir. Unfortunately much of his collection was destroyed in a fire. On his return, he transferred himself to the Iron Works and contributed a number of important papers on this topic to the Irish Naturalist and the J of Botany (1892). Published "The plants of Westmeath" in 1894. The Herbarium was dedicated to him by J.G. Baker. His herbarium is now housed in the National Museum in Dublin.' To the right is a map of Ireland with a red dot in the west. Below the bio is a specimen sheet titled 'WESTMEATH PLANTS. By E. F. and W. B. LEVINGE.' The sheet contains a typed paragraph about the plants, a handwritten note '1364 Ex herb. J.H.L. Levinge', a stamp 'Herbarium of the National Museum Dublin', and a handwritten note 'via Ballymynham. Co. Clare. Ex herb. J.H.L. Levinge May 18. 1892'. At the bottom of the screenshot is the URL <http://herbariaunited.org/atHome/>.

In this example, people transcribing faunal specimen sheets started to realise that they were seeing the same handwriting on different cards and started to wonder about the people behind the collections. They collected examples, started to see relationships, compile biographies. They joined because they were interested in science, but then become interested in history.

http://herbariaunited.org/wiki/Harry_Corbyn_Levinge or
http://herbariaunited.org/wiki/Augustin_Ley



<https://www.flickr.com/photos/stateibraryofnsw/3072281873>

Challenges for museums and libraries

Crowdsourcing as hosting a party



Planning a crowdsourcing project is like planning a party. You think about who you want to invite, create nice invitations, do the work to tidy up so it looks nice when they arrive... When people arrive, you stay in the room with them (you don't lock yourself in the bathroom). Being a host takes time. You have to put time and energy into community interaction and marketing.

Prepare for window shopping

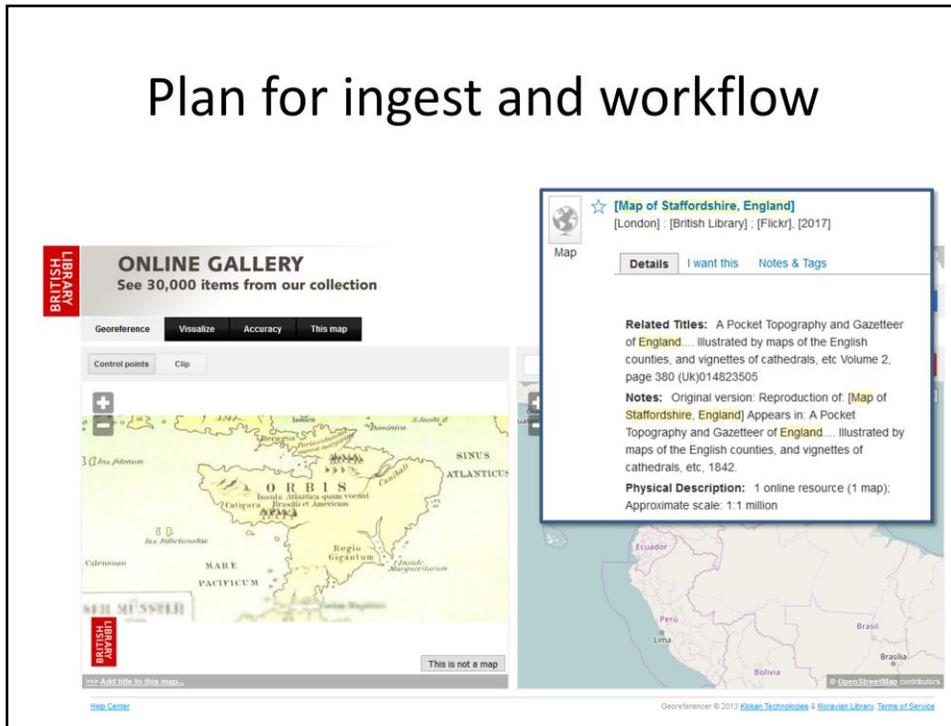


Design is important because there lots of projects around; people will compare yours to others they're tried. If they don't like what they see, they'll walk on.



If a project is going to add to or change their roles, how can you support staff members in this? Do they have the skills to manage technical aspects of workflow or social aspects of community interaction?

Plan for ingest and workflow



Because it wasn't planned from the start, it took a long time to get data from the British Library Georeferencer project into the catalogue as MARC (library format) records. Conversely, changes to Trove text are 'live' immediately.

Getting material from crowdsourcing project into collections management systems, finding aids, etc, is often forgotten or delayed because it's quite hard, but you should plan technical workflow from the start.

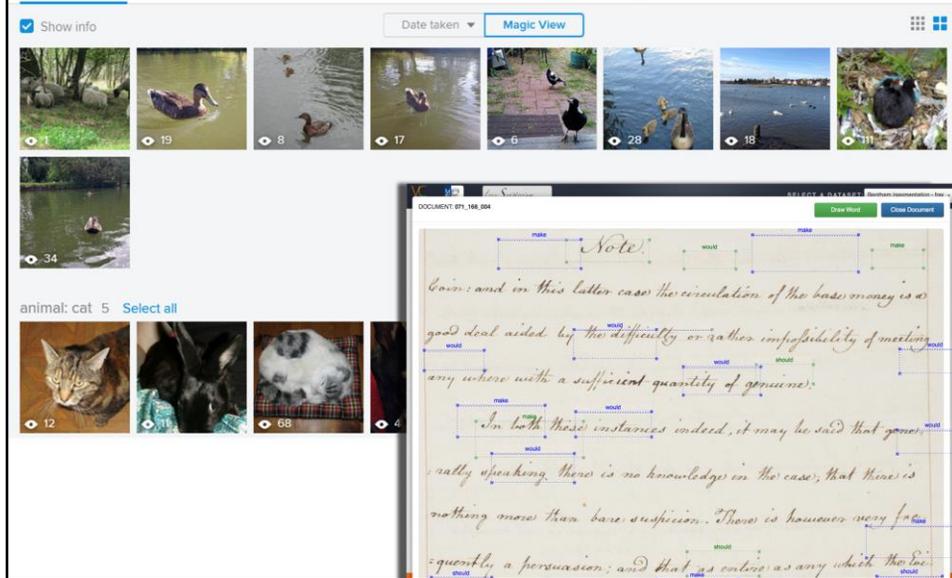
Looking to the future

The growth of platforms

The screenshot shows a web interface for creating a project. At the top, the title "The growth of platforms" is displayed. Below it, the page is titled "PROJECT #829" and "Project details". A sidebar on the left contains navigation links: Research, Results, FAQ, Education, Collaborators, Media, Visibility, and Talk. The main content area is titled "Input the basic information about your project, and set up its home page." It features an "Avatar" section with a placeholder "Drop an avatar image here" and a "NAME" section with a text input field containing "Untitled project 2015-09-09T09:27:39.399Z". Below the name field, there is explanatory text: "The project name is the first thing people will see about the project, and it will show up in the project URL. Try to keep it short and sweet. Your project's URL is </projects/usernameilneverremember/untitled-project-2015-09-09t09-27-39-dot-399z>". A "DESCRIPTION" section has a text input field with "Description of project" and a note: "This should be a one-line call to action for your project that displays on your". At the bottom of the form, there are navigation links for "crowdcrafting", "DISCOVER", and "ABOUT", a user profile indicator "mia" with a dropdown arrow and a "M" icon, and a "CREATE YOUR PROJECT" button. Below the form, there is a banner for "historypin" with the text "Create a Project and get other volunteers to help you analyse your data." and a "CREATE A PROJECT" button. At the bottom left of the banner, there is a logo for "FROM THE PAGE".

Opportunity – easier than ever before to try it out! Less time thinking about tech; more time thinking about people.

Machine learning integrated with crowdsourcing



On the right, a project using crowdsourcing as input into process of teaching computers to read handwriting. In the background, Flickr learnt how to label animals from tags added by people.

谢谢 Questions?

Dr. Mia Ridge, @mia_out
Digital Curator, British Library

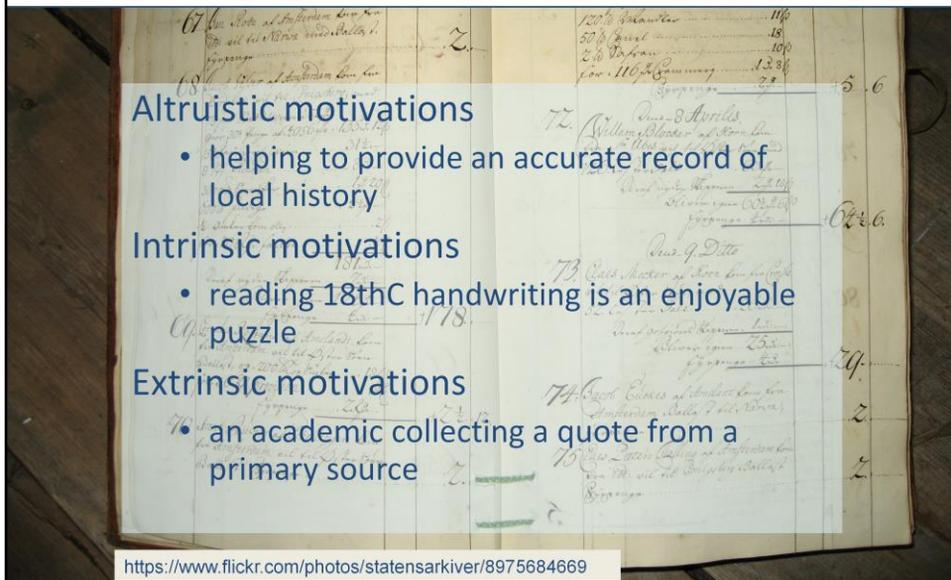
Chinese Association of Museums
Taipei, Taiwan, August 2017

Communities of practice as 'social learning systems'

Old Weather: Whaling		
 Learning the Ropes & FAQ Introduction to marking and transcribing our whaling ship logs	259 Posts 7 Topics	Last post by Randi in Re: Report technical pro... on October 21, 2015, 07:56:32 pm
 Whalers' Shipyard Talk about your ship and crew, and ask ship-specific questions	145 Posts 38 Topics	Last post by Janet Jaguar in Re: Milo - Discussion on October 21, 2015, 03:58:01 pm
Questions and Answers		
 The Dockyard Ship information and transcribing assistance Moderators: Hanibal94 , propriome , AvastMH	11537 Posts 91 Topics	Last post by Hanibal94 in Re: Jamestown (1844) -- ... on October 27, 2015, 01:40:54 pm
 Handwriting Help Finding the handwriting hard to read?	2110 Posts 151 Topics	Last post by Randi in Re: Received on board on... on October 13, 2015, 12:20:05 pm
 The Logs and FAQ General questions and observations Moderator: Kathy	3116 Posts 159 Topics	Last post by Pommy Stuart in Re: Barometers, Instrume... on October 22, 2015, 09:36:16 pm
Library		
 Reference Desk Find useful links and interesting information here	202 Posts 26 Topics	Last post by Janet Jaguar in Re: Sea Ice Types on September 30, 2015, 11:49:57 pm
 Geographical Help Tools for finding the names of towns, bays, points, islands, lighthouses and other landmarks Moderators: propriome , AvastMH	629 Posts 26 Topics	Last post by propriome in Re: Canadian Place Names... on October 26, 2015, 01:26:41 pm
 What Does THAT Mean? Get help with specific text	1310 Posts 137 Topics	Last post by Janet Jaguar in Re: Terms found in whali... on October 01, 2015, 12:23:39 am

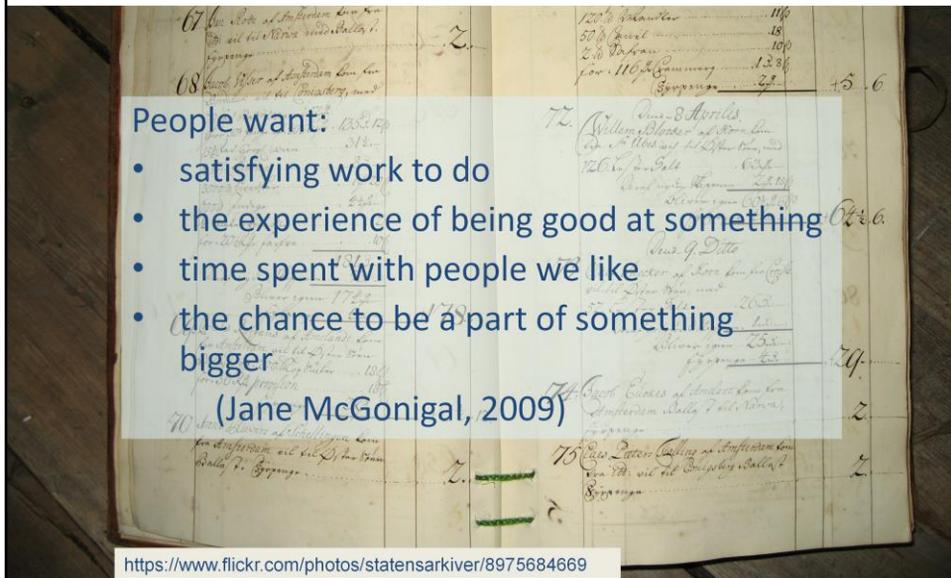
At their best, participant forums become communities of practice, which in turn support learning. In CoPs, newcomers 'learn and acquire knowledge through participating in everyday activity with colleagues'. Online forums support many of the activities typical of communities of practice, including problem solving, making and answering requests for information, coordinating activities and undertaking documentation projects. Online communities of practice develop 'a shared repertoire of resources' including 'experiences, stories, tools, ways of addressing recurring problems'. While the original theoretical work on communities of practice involved in-person discussion, online communication, including social media, forums and discussion lists, similarly show many traces of the development of shared practices.

Understand why people participate



My division of motivations; others may be different.

Use motivations as design guidelines



People want:

- satisfying work to do
- the experience of being good at something
- time spent with people we like
- the chance to be a part of something bigger

(Jane McGonigal, 2009)

<https://www.flickr.com/photos/statensarkiver/8975684669>

Another way of looking at it.. Source: <http://www.aamus.org/resources/publications/museum-magazine/museums-as-happiness-engineers> and <http://www.youtube.com/watch?v=zJ9j7klZuoQ&feature=plcp>

Design is part of recruitment

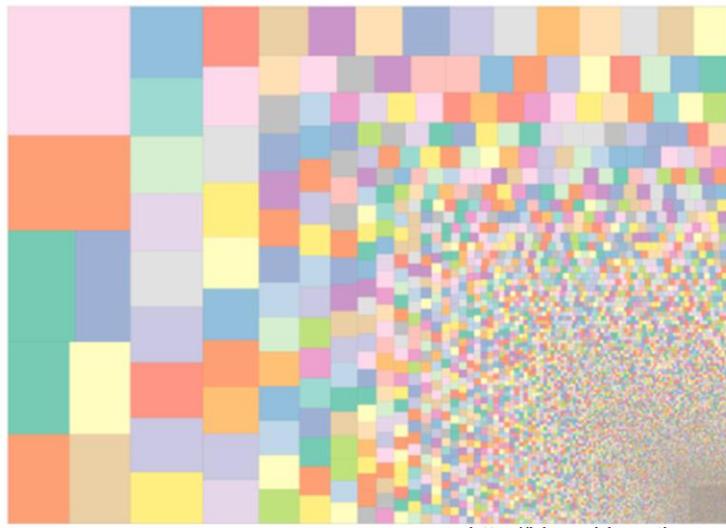


Design really matters, but museums aren't always equipped to work on this.

Finnish example! Easy-to-learn game-play; Simple controls; 'Forgiving' game-play with low risk of failure; Carefully managed complexity levels with a shallow learning curve, guidance through early levels, and inclusive, accessible themes; Sense of rapid progress and achievement = flow!

Build any tests for skill or experience requirements into the interface; Build tutorials for new skills into application at the point where its needed; provide good feedback on actions

Design for casual and super-contributors

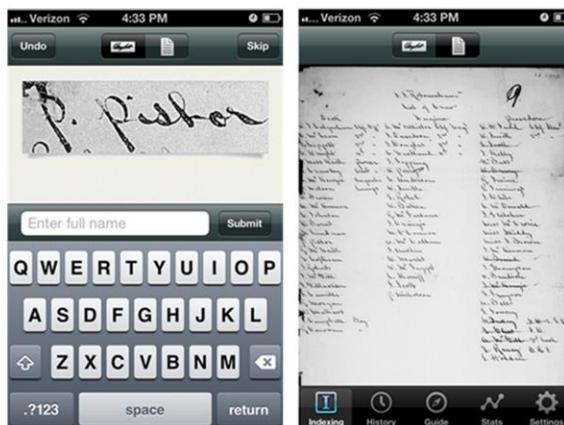


<http://blog.oldweather.org>

Designing for a range of users can be difficult. Some people do a lot of the work, and a lot of people do some of the work. This represents all 16,400 people who have transcribed at least one page for Old Weather (back in 2012)

Source: <http://blog.oldweather.org/2012/09/05/theres-a-green-one-and-a-pink-one-and-a-blue-one-and-a-yellow-one/>

Going mobile



<https://tech.lds.org/blog/455-new-familysearch-indexing-app-now-available>

Transcribing information ('indexing') 'The FamilySearch Indexing app simplifies indexing by allowing you to transcribe individual names, or “snippets,” on your mobile device instead of downloading larger batches of names that must all be transcribed as part of a group. (You also have the option to view the entire document so you can see the name in context.) You can set a difficulty level and skip snippets that are too hard to read. ...To increase indexing accuracy, the app requires snippets to have a 75% consensus to be approved, at least for the app's initial pilot period. ... That number may be adjusted later as genealogists determine the most effective confirmation rate.' Source: <https://tech.lds.org/blog/455-new-familysearch-indexing-app-now-available>

'no plan survives contact with the crowd'

Home About Projects Statistics Data Blog

LibCrowds Community [LOG IN](#)

Implemented ideas

■ Ideas

 alexmendes 2 Jun 10 [Recording Shelf Marks](#)

Here we will record all of your ideas for the platform that been accepted and successfully implemented. [Adding a comment to a task](#)

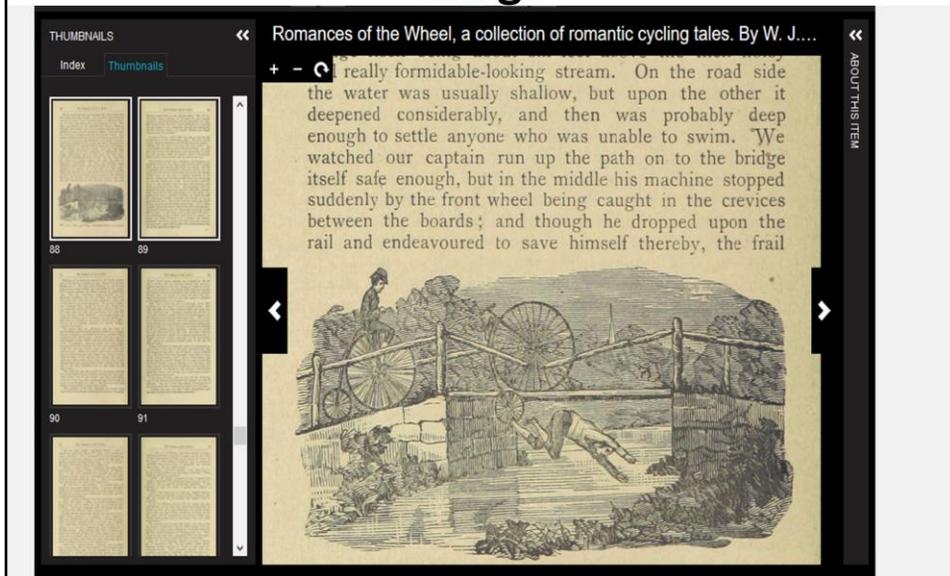
10/06/2015
[Adding comments to a task](#)
 @Adi
 A comments field has been added to the Convert-a-Card projects, allowing contributors to provide additional information that might help ensure a successful and correct match.

31/07/2015
[Draggable cards](#)
 @billmccloy
 All cards seen within the crowdsourcing projects can now be dragged to anywhere on the screen. This provides more flexibility, helping people who wish to organise their screen in different ways and potentially make the transcription process easier.

1 Like 

Projects change once a community finds them. Allow time to update after launch as things will need to be tweaked and participants often have good ideas. Online projects aren't like exhibitions – you need to keep working on them after launch.

From 'deliverable unit' to addressable image



The library's spent over a century honing the art of delivering the physical thing that someone requests to the reading room. This 'deliverable unit' seems obvious if you've ordered a book, but it's less so if you've ordered a manuscript or archive document. You may get a box of documents or a bound volume of manuscripts. Often, the catalogue only contains records for the deliverable item - the book, the box, the volume; not the document, the chapter, the page.

An image like this could have a life of its own on Flickr Commons, where it could be tagged with new terms and links. But it's not easy to honour that work by adding all that extra information back into the catalogue. There's a mismatch in granularity that'll take some time to resolve before we can make items addressable at the manuscript, page and image level in order to capture transcriptions, tags and links that enhance the record.