Integrating Crowdsourcing Projects for Archives Using IIIF

Mark A. Matienzo / @anarchivist
Stanford Libraries
3 December 2018 / #iiif

On behalf of and based on a presentation by:
Josh Schneider, Assistant University Archivist, Stanford Archives / @StanfordArchive
Sara Carlstead Brumfield and Ben Brumfield, FromThePage / @_FromThe_Page
Alumni: Share your student materials with the University Archives!

The Stanford Alumni Legacy Project focuses on collecting, preserving, and providing long-term access to student materials created by Stanford alumni during their time on the farm. The Archives is particularly interested in letters, email, texts, photographs, audio and video recordings, posters, websites, flyers, and records of student organizations. These collected materials are used to support teaching, learning, and research at Stanford and beyond.

We are especially interested in acquiring materials relating to Stanford women, individuals of color, activists, and members of the LGBTQIA community.

You can donate your materials to the Archives or lend them for selective scanning. The Archives will also accept digital content, which will be preserved in the Stanford Digital Repository and made available online. All donations and scans of materials will be preserved and made available for teaching, learning and research. To help guide alumni, the Archives has created a handy list of materials they typically collect.

You can forward any memorable email messages from your time at Stanford to yourstanfordlegacy@stanford.edu, and any memorable texts to (415) 952-7241.

You can also leave a brief audio recording detailing a story or memory from your time at Stanford by leaving an audio message on the University Archives' Stanford Stories line: (415) 952-7241. Please clearly state your full name, Stanford affiliation, class year, and story or...
You can see all page revisions and compare the changes made in each revision. Left column shows the page title and description in the selected revision, right column shows what has been changed. Unchanged text is highlighted in white, deleted text highlighted in red, and inserted text is highlighted in green.

**gkazeber** at Apr 08, 2018 12:29 AM

<table>
<thead>
<tr>
<th>Revision</th>
<th>Date</th>
<th>User</th>
<th>Compare with</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>07, 2018</td>
<td>gkazeber</td>
<td>Oct 18, 2017 - caroleholmson</td>
</tr>
<tr>
<td>Journal of Light-house Station at Cape Foulweather Oregon</td>
<td>08, 2018</td>
<td>Friends of Ya...</td>
<td>13, 2017 - caroleholmson</td>
</tr>
<tr>
<td>June 1891</td>
<td>07, 2018</td>
<td>gkazeber</td>
<td></td>
</tr>
<tr>
<td>22 General duties for the day - Keeper went to Newport for recreation - Left the station at 7:30 A.M.</td>
<td>08, 2017</td>
<td>Friends of Ya...</td>
<td>13, 2017 - caroleholmson</td>
</tr>
<tr>
<td>13</td>
<td>07, 2018</td>
<td>gkazeber</td>
<td>Oct 18, 2017 - caroleholmson</td>
</tr>
<tr>
<td>June 1891</td>
<td>07, 2018</td>
<td>gkazeber</td>
<td></td>
</tr>
</tbody>
</table>
Integration challenges

web → email → staff → CDM → web

(people-ware)

Credit: Jen Wolfe via Shawn Averkamp, [http://ir.uiowa.edu/cgi/viewcontent.cgi?article=1217&amp;context=lib_pubs](http://ir.uiowa.edu/cgi/viewcontent.cgi?article=1217&amp;context=lib_pubs)
We're celebrating Black History Month 2018 in a big way. Transcribe with us and help preserve Black history and culture.

Featured

Celebrating Black History Month
Harriet (Hallie) Sophia Hyde Letters

Collection
Stanford student Letters and Memoirs, 1871-1968

Access conditions

Use and Reproduction
The materials are open for research use and may be used freely for non-commercial purposes with an attribution. For commercial permission requests, please contact the Stanford University Archives (university.archives@stanford.edu).

Copyright
pretty good but-not-wholesome. We always have mush for breakfast and hot-breads - little biscuit rolls or coffee cake, and toast = The meal is as a rule chops or [shank?] and on Sunday ham + eggs cooked the way you wish - we have chocolate Sunday also For lunch we have either baked or boiled potatoes, huge ones milk is always on the table ([blue?]), cold meal -- usually -- and (?) For dinner soup, mashed potatoes - lamb, roast-beef or something else- one vegetable - and desert - which defies analysis On Sunday we have ice cream. But the way the things are cooked don't compare with Jim's delicious wholesome things There are six long tables - longer than ours at home - And

L.S.J.U. November 13, 1894
My dear Mamma - It made me laugh to read your remark about the marmalade If you are here you wouldn't wonder of be suprised that it goes - Several hungry girls can easily make way with a quart of marmalade for a [week?] or so. We have and jam or jelly or pickles, oh do I want a bottle of Jim's delicious crisp pickles It distrested me to learn that you had sent the lamp you use - I thought it was one you had where is you [fig order?] We need a lamp - because when we are out - (?) we have to [stress?] in the dark but I will bring
pretty good but not wholesome.
We always have mush for breakfast
and hot-breeds - with biscuits
rolls or coffee cake, and toast = The
meal is as a rule chops or steak
and on Sunday ham + eggs
cooked the way you wish - we have
chocolate Sunday also
For lunch we have either baked
or boiled potatoes, huge ones -
milk is always on the table ([blue?]!),
cold meal - usually -- and (?)
For dinner soup, mashed potatoes
lamb, roast-beef or something
else- one vegetable - and
desert - which defies analysis -
On Sunday we have ice cream.
But the way the things are
cooked don't compare with
Jim's delicious wholesome things -
There are six long tables - longer ours at home - And
Any IIIF Enabled Repository

Any IIIF Image + Annotation Viewer

FROM THE PAGE
On-Demand Integration
Repository Integration
“Mine”

FromThePage Contributions API

iiif/contributions/stanford.edu

Returns IIIF Collection Manifest
“Recent”

FromThePage Contributions API

contributions/stanford.edu/20180521
contributions/stanford.edu/20180521/20180525
Stanford Student Letters and Memoirs

Export Subject Index
Click the button to export subject indexes from the entire collection in a single CSV file. Please note that the export process could take some time to complete, so please wait until you get the File Save dialog.

Export All Works
Click the button to export all works from the collection into a zip file containing each work as an XHTML file with transcripts, user comments, subject articles, and internal HREFs linking subjects and pages.

Export Individual Works
You can choose to export individual works in two different file formats. XHTML exports a work as a single-page XHTML file with transcripts, user comments, subject articles, and internal HREFs linking subjects and pages. TEI exports a work as a P5-compliant TEI-XML document.

<table>
<thead>
<tr>
<th>Work Title</th>
<th>Pages</th>
<th>Indexed</th>
<th>Transcribed</th>
<th>Review</th>
<th>Progress</th>
<th>Export As</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Semester at Stanford</td>
<td>118 pages</td>
<td>0%</td>
<td>60%</td>
<td>40%</td>
<td></td>
<td>XHTML</td>
</tr>
<tr>
<td>Allbach correspondence</td>
<td>213 pages</td>
<td>0%</td>
<td>2%</td>
<td>0%</td>
<td></td>
<td>XHTML</td>
</tr>
<tr>
<td>Bartlett letters</td>
<td>84 pages</td>
<td>0%</td>
<td>1%</td>
<td>1%</td>
<td></td>
<td>XHTML</td>
</tr>
<tr>
<td>Barzen letters</td>
<td>13 pages</td>
<td>0%</td>
<td>0%</td>
<td>100%</td>
<td></td>
<td>XHTML</td>
</tr>
<tr>
<td>Birkhofer letters</td>
<td>4 pages</td>
<td>0%</td>
<td>100%</td>
<td>0%</td>
<td></td>
<td>XHTML</td>
</tr>
</tbody>
</table>
## FromThePage Status Attributes

<table>
<thead>
<tr>
<th>FromThePage Status Attribute</th>
<th>Translation Status Attribute</th>
</tr>
</thead>
<tbody>
<tr>
<td>pctComplete</td>
<td>pctTranslationComplete</td>
</tr>
<tr>
<td>pctTranscribed</td>
<td>pctTranslated</td>
</tr>
<tr>
<td>pctOcrCorrected</td>
<td>pctTranslationNeedsReview</td>
</tr>
<tr>
<td>pctIndexed</td>
<td>pctTranslationIndexed</td>
</tr>
<tr>
<td>pctMarkedBlank</td>
<td>pctTranslationMarkedBlank</td>
</tr>
<tr>
<td>pctNeedsReview</td>
<td></td>
</tr>
</tbody>
</table>
Letter from Roger Casement (U.S. Consulate General, Munich) to Captain Hans Boehm, thanking him for his letter to the messenger going to Ireland.

metadata:

- label: dc:source
  value: https://data.ucd.ie/api/img/manifets/ivrla:7204

service:

- @context: http://www.fromthepage.org/jsonld/1/context.json
- @id: https://fromthepage.com/iiif/925/status
- label: Work Status
- pctComplete: 100.0
- pctTranscribed: 100.0
- pctOcrCorrected: 0.0
- pctIndexed: 100.0
- pctMarkedBlank: 0.0
- pctNeedsReview: 0.0
- pctTranslationComplete: 0.0
- pctTranslated: 0.0
- pctTranslationNeedsReview: 0.0
- pctTranslationIndexed: 0.0
- pctTranslationMarkedBlank: 0.0
How the workflow might look

- Query contributions API for all our documents in *FromThePage*; returns IIIF collection
- For each item in the collection
  - Request status data from the API
  - Parse response and evaluate status metrics
  - If metrics indicate completeness by our standards
    - Request IIIF manifest from *FromThePage*
    - Capture all page-level transcribed text from annotations referenced in the manifest, process, ingest to repository
    - Capture TEI document referenced in manifest, process, ingest to repository
pretty good but not wholesome. We always have mush for breakfast and hot-breads – little biscuit rolls or coffee cake, and toast. The meal is as a rule chops or shank and on Sunday ham + eggs cooked the way you wish – we have chocolate Sunday also. For lunch we have either baked or boiled potatoes, and on Sunday – prime lamb.
f. 195: Cambridge, Corpus Christi College, MS 053: The Peterborough Psalter and Bestiary

Lamb, Goat, Boar, Bullock, Ox, Camel: two men load it
Stanford Letters Project

Help us transcribe Stanford history! These letters, diaries, and other materials help tell the story of Stanford.
Owner Dashboard

Small Organization account • Current subscription expires Aug 30, 2018 • Total pages: 20886

18 Collections
1,139 Works

Start A Project Your Works Summary

1,139 Works
20,886 Pages
117 Notes
148 Subjects
428 References
123 Collaborators

4,963 Pages transcribed
2,190 Page edits
148 Pages indexed
0 Pages translated
0 OCR Corrections

Last 7 Days Statistics

Collaborators  Pages transcribed  Page edits  Pages indexed  References  New subjects  Note  Pages translated  OCR Corrections

7  31  17  0  0  0  1  0  0
April 20, 1917

At 7:45 P.M. the meeting was called to order by Prof. Law. Due to unavoidable causes, the regular meeting was two days later. The absentees were Murn Bush & Hall.

At 7:30 P.M. the regular meeting of the month was called to order by Prof. Law. The absentees were Murns. Bush & Hall.

The following motion was passed:
(Motion made by Mr. Tung): Our social gathering in honor of the C.S. Club of U.C. be called off on account of unavoidable causes and difficulties. This was seconded by Mr. Tse.

(Motion made by Mr. Kwong): our social gathering in honor of the C.S. Club of U.C. be called off on account of unavoidable causes and difficulties. This was seconded by Mr. Tse.

(Motion made by Mr. Kwong, seconded by Mr. Lee): that our club shall contribute to the Stanford University Library the Chinese

May 4, 1917

Having no other business, the meeting was adjourned.

at 9:35.

Respectfully,

[Signature]

Secretary.
play turns to work
busy getting business planned
Seattle friends
movie w/ Jorge
stability vs capriciousness

Because I'm sure we will see each other again. I will feel comfortable then, if you let you understand I have (it's not heavy, so I'm simply more of a non-competitive person) I'd like to share what's on my mind. Of course the...
Dear Mother,

Thank you very much for mailing my books so promptly. I am happy to report that my books are purchased and things are no doubt successful. Don't worry about sending me money, I do quite well.

Your flu bug has progressed though. I'm sorry you felt so badly and that the kids in my letter bothered you so much. But the way I feel is that no matter how badly you feel, you're just a very important person, so don't worry about it. I hope you feel better soon.

Love,
[Signature]

P.S. Please send me the address for my brother's new job. He's moved to Washington, D.C.
What’s next?

• Getting stuff out; how to provide access
• Dealing with externally sourced content
• Provenance of crowdsourcing and recognition of labor
• Crowdsourcing for photographs
• AI for the other stuff
Questions?

Mark Matienzo
matienzo@stanford.edu / @anarchivist

Josh Schneider
jschne@stanford.edu / @StanfordArchive

Sara Carlstead Brumfield
saracarl@fromthepage.com
@_FromThePage_

Ben Brumfield
benwbrum@gmail.com / @benwbrum