



CROWDSOURCING: Many Hands Get Digitized Records Indexed!

Diane L. Richard looks at why crowdsourcing has become such a valuable benefit to the genealogy community

The last few years has seen a plethora of “crowdsourced” projects benefit genealogists.

First – what does crowdsource mean?

crowd·source

verb [used with or without object], *crowdsourced*, *crowdsourcing*.

1. to utilize (labor, information, etc.) contributed by the general public to (a project), often via the Internet and without compensation: The team's use of Facebook to crowdsource accurate scientific data allowed the project to be completed on time.

Or as the British Dictionary states ...

1. to outsource work to an unspecified group of people, typically by making an appeal to the general public on the internet.

Source: www.dictionary.com

Basically, the idea is to get a lot of people to each do a little (though, no-one will complain if you happen to do a lot!). This means of involving people normally requires little time investment by the volunteer and little to no training and so individuals can “dive in” on a whim and whenever they want (the web never shuts down). Beware though, many

individuals who contribute to crowdsource projects do become hooked!

Second – how have genealogists benefited? Many are familiar with some of the very large crowdsource such as the 1940 census project and Freedmen's Bureau Records and the annual FamilySearch indexathons, and who hasn't benefitted from those? There are many smaller, or shall we say narrower, in scope projects that are just as valuable to our research and which I know have benefitted my research.

Third – now that I've teased you a bit about these projects, let's take a stroll through some of these projects and learn about them and whether they are still active or not. Listed on the following pages in table form are most of the crowdsourced projects that I am aware of and I probably do know of others that I use constantly, and which are not mentioned here; no slight intended. I am also sure that there are other projects that I am not aware of. Please drop me an email (see contact info at end) about the projects you know of. Whether active or not, they could become the subject of a future NetNote or a follow-on to this article. Any inaccuracies are mine; the list was created using only Internet available resources.

Note that these are not all “just” genealogical in nature. Given the importance of historical context to the research we do into our ancestry, several listed projects provide just that kind of contextual information invaluable to ensure that we accurately interpret and understand the records we do find.

These are presented in alphabetical order (based on either project title or institution depending on what seemed more applicable). To provide some additional context, geographic locales have been put in bold.

As you can see, there are quite a few projects covering a variety of records types and time periods all of which are using crowdsourcing as the means of expeditiously getting these materials indexed and/or corrected to increase researcher access. See the following pages for table data.

Page from the Smithsonian Digital Volunteers: Transcription Center showing the progress on various Freedmen's Bureau transcription projects.

Title and/or Institution	URL	Description	Status
1940 Census	www.familysearch.org/blog/en/1940-census	A joint initiative between the National Archives and Records Administration (NARA), Archives.com, FamilySearch.org, findmypast.com, and other leading genealogy organizations to create a free, high-quality, searchable database of the 1940 US census records	Completed
Ancestry – World Archives Proj.	www.ancestry.com/community/awap	Typically has multiple projects from around the world in process. [Started 2008]	Active project. See URL for participation details.
Anti-Slavery Manuscripts – Boston Public Library [Mass.]	www.antislaverymanuscripts.org	Turn collection of handwritten correspondence between anti-slavery activists in the 19th century into texts that can be more easily read and researched by students, teachers, historians, and big data applications.	Active project. See URL for participation details.
British Library	www.libcrowds.com	A platform dedicated to the hosting of experimental crowdsourcing projects aimed at enhancing access to British Library collections. An initial project, Convert-a-Card, sought to retro-conversion printed card catalogues into electronic records.	Active project. See URL for participation details.
Decoding the Civil War	www.zooniverse.org/projects/zooniverse/decoding-the-civil-war	Provide open source access to a large, rare collection of 15,971 Civil War telegrams in The Thomas T. Eckert Papers that are owned by The Huntington Library [California]	Phase 1 Completed, Phase 2 Coming Soon.
Europeana Transcription 1914-1918	http://transcribathon.com/en	Online crowdsourcing initiative for the transcription and annotation of this unique, unpublished material. The aim is to create a vast and fully digital record of the written stories in Europeana 1914-1918.	Active project. See URL for participation details.
FamilySearch	www.familysearch.org/indexing	Typically has 100 or more projects from around the world in process. [Started 2006] Familysearch has previously hosted worldwide indexing events (last one was 20-22 October 2017)	Active project. See URL for participation details.
Discover Freedmen	www.discoverfreedmen.org	A set of partnerships to facilitate the indexing of the Freedmen’s Bureau records in anticipation of the opening of the National Museum of African American History and Culture. The project began on Juneteenth (June 19) 2015 and was completed on June 20, 2016. Over twenty-five thousand volunteers helped.	Completed



Title and/or Institution	URL	Description	Status
Immigrant Ships Transcriber Guild (ITSG)	www.immigrantships.net	Provide assistance in transcribing and/or acquiring immigrant ships passenger lists.	Active project. See URL for participation details.
Letters of 1916 [Ireland]	Letters of 1916, http://letters1916.maynoothuniversity.ie	Begun in September 2013, this digital collection includes letters held at institutions in Ireland and abroad alongside those in private collections. There are hundreds of letters connecting thousands of lives commenting a myriad of topics including the Easter Rising, literature and art, the Great War, politics, business, and ordinary life. [in 2016, project collection period extended to 1923]	Active project. See URL for participation details.
Archives of Latvian Folklore	http://lv100.garamantas.lv/en [English language interface]	Summary about project in English can be found here, http://dig-hum-nord.eu/projects/crowdsourcing-platform-for-folklore-manuscript-transcription	Active project. See URL for participation details.
Library of Congress – Flickr Commons Project	www.flickr.com/photos/Library_of_Congress	A project started in 2008 to identify and explain historical photos by tagging them. Collection now includes more than 30,000 images	Active project. See an image you know something about, add a tag, note or comment.
Library and Archives of Canada	http://t8.ourdigitalworld.org/en	Lady Macdonald's Handwritten Diary & The Coltman Report, 1818 – Inquiry into the offences committed at the Battle of Seven Oaks	Completed
Library of Congress – Beyond Words	http://beyondwords.labs.loc.gov/#/	Project started in 2017 to identify and improve information about illustrated newspapers in the Chronicling America digital collection.	Active project. See URL for participation details.
Library of Virginia – Making History: Transcribe	www.viriniamemory.com/transcribe	Collaborative online workspace where the public can participate in enhancing access to many amazing collections documenting over 400 years of Virginia history, people, and culture.	Active project. See URL for participation details.
Digital Maine Transcription Project	www.digitalmaine.net/projects	DMTP started in 2016 as a collaborative project of the Maine State Archives and Maine State Library.	Active project. See URL for participation details.
Seeking Michigan – Naturalization Index Project	http://seekingmichigan.org/naturalization	A digitization and indexing project to make naturalization records from nearly 70 Michigan counties freely available online. [In partnership with FamilySearch]	Active project. See URL for participation details.

Title and/or Institution	URL	Description	Status
NARA – Citizen Archivist Dashboard	www.archives.gov/citizen-archivist	To make records more discoverable, you can help crowdsource metadata and information about records through tagging, transcribing and adding comments in the National Archives Catalog.	Active project. See URL for participation details.
National Library of Ireland – Photo Detectives	[see description]	In June 2011, the library joined Flickr Commons and over the course of 6 years, some 34,000 “photo detectives” helped uncover details about the photos in the collection. Read more here, www.atlasobscura.com/articles/photo-detectives-ireland-irish-history-archives-exhibition-crowdsource and www.irishtimes.com/culture/heritage/amateur-sleuthing-helps-solve-the-mysteries-of-old-irish-photos-1.3220079	Completed
Transcribing Modern Manuscripts – Newberry Library [Chicago]	http://publications.newberry.org/digital/mms-transcribe/index	Transcribe letters, diaries, journals, and other material from the Newberry’s Modern Manuscripts Collections, a repository of American manuscripts from the mid-18th through the 20th centuries. The materials comprise a range of items, including letters exchanged by three generations of an immigrant family, the Everetts, and the Civil War diaries of Private Ferdinand Kilbourne, of Covington, Pennsylvania.	Active project. See URL for participation details.
NYPL Labs & related [New York City]	www.nypl.org/collections/labs	Some of the projects created: Direct Me NYC 1940, http://directme.nypl.org Navigating the Green Book, http://publicdomain.nypl.org/greenbook-map Building Inspector, http://buildinginspector.nypl.org Maps by Decade, http://spacetime.nypl.org/maps-by-decade/# Stereogramator, http://stereo.nypl.org NYPL Map Warper, http://maps.nypl.org/warper What’s on the menu?, http://menus.nypl.org	Closed
New York Times – Madison	http://web.archive.org/web/20170915072428/http://madison.nytimes.com:80	The New York Times archives are full of advertisements that give glimpses into daily life and cultural history. Help digitize historic ads by answering simple questions. [article in 2014 about, www.nytimes.com/times-insider/2014/10/16/viewing-old-times-ads-with-a-new-tool-called-madison] [the New York Times Research & Development group is still active, http://nytlabs.com]	Closed



Title and/or Institution	URL	Description	Status
Old Weather	http://whaling.oldweather.org/#	Volunteers assist in exploring, marking, and transcribing ship's logs, largely from the 19th and early 20th centuries.	Active project. See URL for participation details.
Operation War Diary (UK)	www.operationwaridiary.org	The story of the British Army on the Western Front during the First World War is waiting to be discovered in 1.5 million pages of unit war diaries.	Active project. See URL for participation details.
Parochial Archive Project in Rome (Italy)	www.zooniverse.org/projects/vincen-t-/parochial-archive-project-in-rome/classify	You need to transcribe the entire folio. Documents are in French, Italian, and Latin.	Active project. See URL for participation details.
Stanford University Archives – From the Page	www.fromthepage.com/stanforduniversityarchives	A crowdsourcing project to transcribe handwritten letters and documents within its holdings. The project currently features 8 unique collections for users to transcribe: 1906 earthquake and fire, Leland Stanford, Jr. letters, World War I letters, World War II letters, etc.	Active project. See URL for participation details.
State Library of Queensland [Australia]	www.slq.qld.gov.au/about-us/pitch-in/transcribe	Make SLQ collections easier to read and searchable.	Active project. See URL for participation details.
Smithsonian Transcription Center	http://transcription.si.edu	Join 10,748 "volunpeers" to add more to the total 351,626 pages of field notes, diaries, ledgers, logbooks, currency proof sheets, photo albums, manuscripts, biodiversity specimens labels that have been collaboratively transcribed and reviewed since June 2013	Active project. See URL for participation details.
Trove – National Library of Australia	http://trove.nla.gov.au/newspaper	A program whereby individuals can correct OCR created text to reflect what was actually printed in the newspapers (now digitized). Guidelines for participating, http://help.nla.gov.au/trove/digitised-newspapers/text-correction-guidelines	Active project. See URL for participation details.
Washington State Archives	http://scribe.digitarchives.wa.gov	Using Scribe (registration required), you can help transcribe records for public access.	Active project. See URL for participation details.

There can also be crowdsourcing projects where the outcome is not to help index projects and it might be to fund indexing projects and/or solicit

help in other ways. For example, The Federation of Genealogical Societies launched the Preserve the Pensions Project, www.preservethepensions.org [fully

funded – raised more than \$US 3 million] and then, with BillionGraves, launched an initiative to “image all cemetery markers for War of 1812 Participants”, <http://blog.billiongraves.com/2014/07/billiongraves-and-the-federation-of-genealogical-societies-partner-to-image-war-of-1812-participants>. Speaking of cemeteries, Find A Grave, www.findagrave.com, has been around since 1995, relying on contributions from volunteers to document final resting places. Want to contribute? Visit here to learn how, www.findagrave.com/contribute.

Another type of project revolves around funding; not everything can be accomplished by volunteers contributing their valuable time. For example, recently, the State Archives of North Carolina sought to crowdsource funding in support of identifying the contents of a photograph collection, www.indiegogo.com/projects/fund-the-work-of-raleigh-s-photo-history-detective-photography# [collection on Flickr, www.flickr.com/photos/north-carolina-state-archives/albums/72157681748351705, and finding aid, http://ead.archives.ncdcr.gov/PHC_BARDEN_Albert_Barden_Photos.html].

Last and not least, maybe you or an organization you are associated with would like to set up a crowdsourcing project. I suggest that you Google (or use the search engine of your choice) on the topic of crowdsourcing. A free resource available for your use is Crowd Sourced Indexing (CSI), <http://csindexing.com>, created by Banai Lynn Feldstein – “Crowd Sourced Indexing is a program to facilitate crowd-sourced projects by ... smaller organizations, usually

comprised entirely of volunteers, [which] have no access to [custom] programs ... CSI is easier for indexers. Instead of juggling multiple windows, to view an image and a spreadsheet program at the same time, everything is presented in one browser window. Volunteers can log in and receive the next image whenever they feel like indexing. There is no downloading of the images, sending in their transcriptions, and waiting for another batch. When they have indexed a page, they can decide if they want to continue to do more immediately or continue another time.” The main page lists projects being currently worked on and you can sign up and participate.

Alternately, one can also use Dropbox, Google Drive, and a myriad of other techniques to establish crowdsourcing projects. I regularly engage volunteer transcribers, indexers, etc., for genealogy journal content using just email and Dropbox. With all the options now available to us in terms of sharing material, crowdsourcing is NOT just for the more formal projects by large entities, as presented, and can be successfully utilized by individuals and smaller organizations. ©

DIANE L. RICHARD has been doing genealogy research since 1987. She is currently editor of North Carolina Genealogical Society Journal and Wake Treasures (journal of the Wake County Genealogical Society) and a professional genealogy and family historian researcher, speaker, and writer. She can be found online at www.mosaicrpm.com and www.tarheeldiscoveries.com.

To learn about more crowdsourced or participatory projects, check out these resources:

ConferenceKeeper.org – Genealogy Volunteer Activities, <http://conferencekeeper.org/volunteer> [updated regularly]

Crowdsourcing and Participatory Archives Known Projects [dated May 2014], <http://repository.asu.edu/attachments/135632/content/Known%20Projects%20Crowdsourcing%20and%20Participatory%20Archives.pdf>

List of crowdsourcing projects [Wikipedia], http://en.wikipedia.org/wiki/List_of_crowdsourcing_projects – these are NOT just genealogy projects and if you have the patience, it lists both complete and active projects.

Cultural Institutions Embrace Crowdsourcing [The Signal – Library of Congress Blog], <http://blogs.loc.gov/thesignal/2015/09/cultural-institutions-embrace-crowdsourcing> [dated September 2015]

Crowdsourcing Digital Public History [The American Historian], <http://tah.oah.org/content/crowdsourcing-digital-public-history> [not dated]

All of this means that there are many ways that you can HELP with these types of projects while you also BENEFIT from what they produce – increased access to genealogically relevant materials.