

UC Berkeley

UC Berkeley Previously Published Works

Title

From collection curation to knowledge creation: Exploring new roles of academic librarians in digital humanities research

Permalink

<https://escholarship.org/uc/item/5dn0r0kp>

Journal

The Journal of Academic Librarianship, 47(2)

ISSN

0099-1333

Authors

Zhang, Ying
Xue, Susan
Xue, Zhaohui

Publication Date

2021-03-01

DOI

10.1016/j.acalib.2021.102324

Peer reviewed

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

The Journal of Academic Librarianship

journal homepage: www.elsevier.com/locate/jacalib

From collection curation to knowledge creation: Exploring new roles of academic librarians in digital humanities research

Ying Zhang^{a,*}, Susan Xue^b, Zhaohui Xue^c^a University of California, Irvine, 466 Langson Library, Irvine, CA 92623, USA^b University of California, Berkeley, C.V. Starr East Asian Library, Berkeley, CA 94720, USA^c Stanford University, East Asian Library, Palo Alto, CA 94305, USA

ARTICLE INFO

Keywords:

Digital humanities
Academic librarianship
Open access
Librarian-faculty collaboration
Global outreach
Metadata
Crowdsourcing translation
Scholarly communication

ABSTRACT

We present a case study of a collaborative digital humanities project that was led by academic librarians and participated by IT specialists and international scholars. Through the project, we produced and published two open-access products – a bilingual dictionary of historical government official titles, and an online system for crowdsourcing translations. We also contributed metadata, from the dictionary, to an existing large digital humanities project. Whereas the dictionary fills a publication gap and a research need, the crowdsourcing system provides a digital research method for collaborative translations of specialized terms. During this process, we explored and undertook a variety of new roles, ranging from project leader, hybrid scholar to content creator and publisher. Our project demonstrates how academic librarians could dive deep in the realm of digital humanities and plunge into the life cycle of scholarly communication.

Introduction

The constant rise of digital humanities (hereafter DH) research calls upon academic librarians to engage in and lead DH programs and projects, and to expand our roles into different stages of scholarly communication, from data collection and analysis, authoring, peer review, publishing to discovery and dissemination. Meanwhile, the advancement of digital technologies provides opportunities for us to step out from unostentatious offices and collections and to explore a new area of facilitation of and participation in knowledge creation.

In this paper, we report our experimental and highly collaborative project, funded by the Andrew Mellon Foundation under the Mellon-CEAL Innovation Program. Under the project, we took the initiative to coordinate collaborations among library professionals, information technologists, expert consultants and scholars from Asia, Europe and North America. Together, we created a bilingual dictionary for the study of China's Ming Dynasty (1368–1644), utilizing a crowd-translation tool that we developed for scholars to participate and contribute at a global-scale. We published both the bilingual dictionary and the crowd-translation system for open-access. Meanwhile, the thousands of official titles published in the dictionary provide a means for creating consistent metadata to be used by other digital projects, including China

Biographical Database or CBDB, for improved discovery and access. Whereas the dictionary fills a publication gap and a DH research need, the crowdsourcing system provides a controlled method for translations of specialized terminology. Our work is a DH research by definition, because it is “concerned with the intersection of computing and the disciplines of the humanities. It is methodological by nature and interdisciplinary in scope” (Kirschenbaum, 2010). We have devised a digital collaboration tool and used it to generate new knowledge for historical research.

Through reporting and evaluating this DH project, we discuss how academic librarians could take advanced roles as proposed by Zhang, Liu, and Matthews (2015), and engage in-depth with scholarly communications as suggested by ACRL (Association of College & Research Libraries, 2003). Under this project, we explored and undertook a variety of new roles, ranging from project leader, hybrid scholar to content creator and publisher. We collected, analyzed and organized a large volume of data, developed a peer review process through crowdsourcing, published two open-access resources, and provided a means for creating consistent metadata for online discoverability. Our research activities are in perfect line with the five steps, outlined in ACRL's scholarly communication lifecycle (Association of College & Research Libraries, n.d.) including research data collecting, creation, peer review,

* Corresponding author.

E-mail addresses: yingz@uci.edu (Y. Zhang), sxue@library.berkeley.edu (S. Xue), zhxue@stanford.edu (Z. Xue).

publication, and discovery.

The rest of the paper is organized in three parts. The first section provides an overview of the project, including its scholarly and innovative background, international collaborative nature, open-access products, as well as post-project activities. The second section discusses the implications of this project on academic librarianship in the digital age. Finally, the [Conclusion](#) section summarizes the significance and future direction of our project.

Project background, method and outcome

Identifying research needs through interaction with scholars

The project idea was formed at a professional meeting back in 2015 where Harvard Professor Peter Bol and Dr. Ying Zhang, the first author of this article, spoke at the same East Asian studies faculty and librarian panel on Vision and Collaboration beyond Local Collections. In reaction to Professor Bol's expressed needs for CBDB, a DH project he leads, Zhang quickly formed a team of Chinese studies librarians in the U.S. to develop a metadata infrastructure for organizing and retrieving government official titles of the Ming Dynasty. Our major tasks were (1) to create a hierarchical tree, describing the entire government system, from imperial court and royal family, central government offices to regional and local offices, (2) to compile as complete as possible a list of bureaucratic posts/official titles and to pair them with corresponding government offices on the hierarchical tree, and (3) to complete English translations of the government offices and official titles, which were not readily available in existing reference sources, mainly Charles Hucker's *A Dictionary of Official Titles in Imperial China* (Hucker, 1985).

These tasks are crucial to CBDB (CBDB, n.d.-a), a massive relational database with biographical information of almost half million historical figures, primarily from the 7th through 19th centuries. Its development relies on a combination of authority files (offices, official titles, places, kinship, social relationship, etc.) and value tables (individuals' names, birth/death years, specific government posts taken, etc.). Although the imperial government system of each dynasty in China has some continuity with its preceding system, it has distinctive features as well. Therefore, authority files need to be constructed separately for each dynasty. Having had related authority files developed for most other dynasties, CBDB was in need of filling in the gap for the Ming Dynasty, which lasted 276 years (1368–1644) and experienced major changes in its political system. Similar to other authority files, the hierarchical structure of Ming government offices and official titles is essential for biographical data retrieval for and access in the relational database. Meanwhile, the standardized English translations of government offices and official titles could save scholars from considerable individual effort trying to devise original English nomenclature, especially since many of them have long been obsolete.

In his letter of support for our grant proposal, Professor Bol wrote, "Librarians, after all, understand the importance of authority files much better than historians generally do... we can only succeed in building a strong cyberinfrastructure for Chinese studies if librarians play a leading role...Scholarship in a digital environment is necessarily collaborative, but by and large the research projects of faculty have not involved robust links to librarians and libraries. The proposed project is a small but vital step towards creating those links and building the cyberinfrastructure we so very much need."

Thereafter, four librarians became collaborators of Professor Bol's DH research team (CBDB, n.d.-b), joining the scholars in the front-line. The librarian team is equipped with basic knowledge and skills that the project requires. In addition to training in librarianship, two team members hold degree in Chinese history, whereas the other two major in political science and information science respectively. Together, we seized the opportunity to embed DH research to librarianship, to experiment collective knowledge creation through crowd-translation, and to build a *Chinese-English Dictionary of Ming Government Official*

Titles to meet a research need and fill a publication gap.

Throughout this project, we received constant support from CBDB, including sharing data sets, and promoting our project on their website. We are happy to see that the hierarchical office tree and official title table developed under this project have been adopted by CBDB. From the time when we announced the project plan and called for contribution at academic listservs and conferences, we already started to receive overwhelmingly optimistic feedback, including "all very exciting", "important contribution", "the willing of a network of librarians to help is most welcome", "should the day ever come when I can do actual work again", etc. Those positive support and feedback not only motivated us to carry out this project, but also laid a foundation of scholars' participation and utilization.

Successful grant applications

The project we undertook is complex and ambitious, for it requires not only a tremendous amount of time, but also a high level of technical and academic expertise, beyond our capability and capacity. Seeking external resources, personnel and financial, was the only path to pursue.

In 2016, we applied and successfully competed for a \$53,810 grant from Andrew Mellon Foundation under CEAL-Mellon Innovation Grant for East Asian Studies Librarians (2016–2017) (Council on East Asian Libraries, 2016). The one-year grant for the project, titled "From Curation of Collection to Creation of Knowledge: Building a Bilingual Dictionary of Ming Government Official Titles through Expert Crowd-Translation," allowed us to hire (1) a programmer to build a crowd-translation system, (2) a research specialist to process data and manage the system, and (3) two historians, who specialize in Ming government system, as expert consultants, to ensure a quality delivery of the bilingual dictionary. The external resources are vital for us to take the lead, collaborating with scholars and IT specialists to produce two open access products: an online platform for crowdsourcing translations of subject-specific terminologies and a *Chinese-English Dictionary of Ming Government Official Titles*.

The success of the CEAL-Mellon grant project attracted a new funding, which allowed us to build a technical infrastructure for sustaining the project. In early 2018, after the librarian team presented the project at a professional meeting, the leading author was invited by the Secretary of the Geiss Hsu Foundation to apply for a new research grant to continue the project. In early 2019, she was awarded a grant of nearly \$36,000 from the Foundation (The Geiss Hsu Foundation, 2019) for a Phase II project (2019–2020), titled "Enhancement and Enrichment of a Chinese-English Dictionary of Ming Government Official Titles". With the new grant, we developed an Application Program Interface (API) to connect UC Irvine's crowd-translation system and Harvard's China Biographical Database, so that any future updates of Ming government official titles in one system would be synchronized in the other. We also enhanced the bilingual dictionary with added Chinese and English indexes and essential addition to and correction of main entries.

Collaboration among librarian, scholar and IT specialist

The crowd-translation project under the CEAL-Mellon Innovation Grant is a highly collaborative project among four groups, including librarians, scholars, expert consultants, and information technologists. Core tasks and activities of each group are as followings.

- (1) The librarian team took the lead. Together with our capable and hard-working contract researcher, we completed a variety of tasks, including historical research mainly on political system in China's Ming Dynasty, information organization, system design, workflow management, outreach, as well as dictionary compilation and publishing. Whereas all tasks are essential to our project, the outreach effort is particularly worth noting for its critical role in making our project widely known and attracting more

contributions from Ming scholars at a global scale. Elaborations on the global outreach activities can be found later in the [Discussion](#) section.

- (2) The expert consultant team, composed of four established Ming historians from Princeton, West Point (United States Military Academy) and Beijing Administrative College, provided invaluable advice and guidance to ensure the right direction of the project and the quality delivery of the dictionary. Among the four expert consultants, two specialize in the Ming government system. In general, the team contributed to the project's success by offering comments and suggestions to the librarian team on the project, by reviewing English translations and making final selections, and by helping to promote the crowd-translation project and the bilingual dictionary.
- (3) Ming scholars around the world were invited and encouraged to make contributions via the crowd-translation system, a virtual collaborative space we created for them. By the official closing date of the project, there had been in total 37 registered and approved international scholars in the system. Considering Ming studies are a highly specialized research field with a relatively small community, this number meets our expectation. The geographical distribution of these scholars are 21 from the United States, 7 from China, 3 from Canada, and 1 each from United Kingdom, France, South Korea, Hong Kong, Thailand and Singapore. Out of the 37 registered scholars, 17 made at least one contribution. The 46% contribution ratio is much higher than Jacob Nielsen's 90-9-1 rule (Nielsen, 2006), which states that, in most online systems, only 10% people ever make contributions, and 90% are "lurkers" who never contribute.

Scholars' contributions are primarily in these six areas –

- submit English translations of official titles at their selection
 - anonymous review others' submissions
 - suggest a peer scholar as the translator for a specific official title
 - comment on Pinyin and/or Chinese titles
 - submit new official titles
 - participate in hot topic discussions through the bulletin board on the platform.
- (4) UCI Libraries' IT unit and our contract programmer from the San Francisco area were the IT backbone of the project. The core tasks of these IT specialists included system development and maintenance, as well as source code upload for open access.

Collaborations among members of these four groups essentially took place in virtual environments, for our collaborators were from different regions of the world. Whereas emails, Skype meetings, Google docs, wiki pages and Slack teams were major channels for librarian team members and IT specialists to communicate with each other and to keep project documentation, the crowd-translation system served as the online collaborative space among scholars, and between scholars and the librarian team.

The system has three types of user account – regular, admin, and super-admin. Once signed-in their regular accounts, scholars could browse/search official titles and submit English translations. They could anonymously review and comment on submissions from others. They could also suggest new official titles and recommend a qualified scholar for a specific official title. Their suggestions and recommendations would be sent to the project leader and research specialist for consideration. On the back end of the system, the project leader and research specialist, who had the super-admin account, could assign an untranslated official title to a registered scholar or send an invitation to a scholar not in the system, requesting contribution. The four expert consultants, each having an admin account, could approve or disapprove translation submissions, especially in case multiple translations

received from the crowd for a single official title. In addition, scholars, once signed-in, could participate in discussions about general topics, which were posted by the librarian team on an interactive bulletin board, such as whether official and quasi-official references to those who passed imperial examinations at different levels should be included in the dictionary.

Multi-pronged quality control

Under this project, we have used multi-pronged quality control to make sure the dictionary meets academic standards. In addition to expert guidance from the four consultants, a three-layer quality control mechanism was built into the crowd-translation system. The three layers include credential authentication, anonymous peer-review, and expert judgment. Anyone, who would like to make contribution, needs to first sign up on the crowd-translation platform. The sign-up form has some required fields, such as institution affiliation and areas of research. Upon receiving a sign-up request, the super-admin reviews and approves (or disapproves) it from the back end. The project team may revoke one's participant status for any inappropriate submission from this person.

Whereas the credential authentication process screens subject knowledge of each potential contributor, the anonymous peer-review mechanism secures the quality of official title entries. In addition to submitting new translations, scholars, once signed-in, may also review and comment on existing submissions from others, on a double-blinded basis. As the ultimate form of peer-review, our Ming historian consultants, who are fluent/native in both Chinese and English languages, approve (or disapprove) translation submissions, using the expert judgment feature on their admin account page.

These built-in quality control functions are essential to maximize benefits of the crowd-translation system, as suggested by Anastasiou and Gupta (2011), and ultimately to build a high-quality dictionary. Having the multi-pronged quality control in place makes our crowd-translation system a reliable digital tool for compiling a bilingual dictionary of any subject that requires in-depth domain knowledge and collective input.

Two open access products

The project has produced two open access (OA) products—a digital tool for crowd-translation and a bilingual dictionary as a reference book. The online system <http://mingofficialtitles.lib.uci.edu/#/>, which librarians designed and IT specialists implemented, served as a virtual community, where scholars around the world worked together to contribute English translations of Ming government official titles that had not been translated in existing publications. On the front end, scholars, once signed in, can browse by hierarchical offices or enter a keyword in the search box for specific official titles to review and translate. They may also just simply pick one from the top of the official title list, sorted by number of English translation titles ascending. Under Profile, a scholar may see all his or her submissions and make edits, if necessary. At a first submission, a scholar's name automatically display on the Contributors page for recognition. On the admin portal, the project team can approve (or disapprove) scholar, add (or archive) official title, approve translation, and export data. The expert consultants can review and approve English translations. Different from many other crowd-translation applications, our system has a built-in, triple quality-control method, allowing for credential authentication, anonymous peer-review, and expert judgment. Available at <https://github.com/UCI-Libraries/Ming-Titles-Dictionary>, the source codes can be downloaded for free and repurposed for compiling bilingual dictionaries of any subject domain, on "a digital platform for research but also for crowd-based participation." (Blanke, Kristel, & Romary, 2017).

Whereas the crowd-translation system serves as the digital research method, the *Chinese-English Dictionary of Ming Government Official Titles* fills a publication gap and "corrects the weaknesses" of library collections (Kennedy, 1983) for expressed needs from Ming scholars and a

well-established DH project – China Biography Database or CBDB. The dictionary was not simply a list of government official titles in alphabetical order. Instead, it is built on the basis of an established government structure. Thousands of official titles in the dictionary are organized and presented in a three-tier hierarchical structure, reflecting the entire government system in the Ming Dynasty. The top tier comprises nine categories of government offices: Imperial Family and Royal Court, Central Governing, Central Administrative Assistance, Legislation and Censorship, Nanjing Capital, Regional and Local Governance, Horse/Salt Business and Maritime Trade, Central and Capital Militaries, as well as Regional Military and Security Units. Under the second and third tiers are dozens, and then hundreds, of cascaded lower divisions, such as ministries, bureaus, commissions and offices. Official titles in the same division are listed in descending order of their ranks and importance. Also included in the dictionary are Civil and Military Honorary Titles, Official and Candidate Titles for the Imperial Examinations, Noble Titles, Vassal Kings, and Consorts and Concubine Titles. Therefore, the dictionary provides a clearly defined governing structure and associated official titles for government offices from central to local.

The dictionary has been uploaded to eScholarship, the University of California's institutional repository, for OA. At its permanent URL <https://escholarship.org/uc/item/2bz3v185>, users can browse and search 3245 Ming government official titles or download the dictionary for offline use. Within its first three years of release, the online dictionary received nearly 12,000 hits and was downloaded almost 4000 times, which is over 300 visits and 100 downloads per month on average. This is a great attention and usage to a specialized dictionary. Feedback from scholars has been overwhelmingly positive – “it is very impressive,” and “very, very helpful.” Furthermore, the government official titles and hierarchical structure for building the dictionary have been adopted by CBDB for enhancing its metadata infrastructure. Our faculty collaborator acknowledged the value of the dictionary as follows.

“(The dictionary) has made an important contribution to the infrastructure for the international study of Ming history. It is more extensive than the set of Ming official titles in Hucker’s dictionary. It provides the structure of the Ming bureaucracy. It provides consistent English translations.”

The bilingual dictionary could benefit not only the DH research of the CBDB team, but people who are interested in different aspects of the Ming Dynasty, including social life, military, economics, religion, education, and literature. Additionally, the English translations of Chinese historic official titles would allow western scholars, who may not necessarily study China, to conduct comparative research of government systems between the East and the West.

We are happy to see many inbound links to the online bilingual dictionary from academic websites, including the China Biographical Database at Harvard, Documentation: Guide to Ming Studies at the University of British Columbia ([University of British Columbia, n.d.](#)), and the Ming History English Translation Project at UCSD ([The Ming History English Translation Project, n.d.](#)). The dictionary has been cited by scholars in their publications as well. For example, Bozhong Li, a chair professor at Hong Kong University of Science and Technology, referred “very occasionally” to the bilingual dictionary “for the English translations of the official titles and positions” in his new book on early modern economy in China ([Li, 2021](#)). According to Lixiang Qian, a Ph.D. candidate at Zhejiang University, all “the Chinese–English translation of Ming government official titles” in his new research paper on DH were from our dictionary ([Qian, 2020](#)). Elke Papelitzky, a postdoc in Belgium, also consulted the dictionary for her research paper on early modern Chinese history ([Papelitzky, 2019](#)).

Discussion

International collaborations among librarians and scholars

Literature review reveals that one common theme across librarians' new roles in DH is collaboration, which is in line with the case study findings of [Golomb, Braunstein, and Hartsell-Gundy, \(2015\)](#). Golomb et al. also suggest that subject librarians need to become collaborators with their colleagues within library and with faculty outside the library. [Kear, Joranson, and Humanities \(2018\)](#) analyze examples of digital humanities collaborations, which librarians have been part of, and recommend proactive ways to build partnership.

Nevertheless, all the DH projects from our literature review were either initiated and completed by librarians themselves or led by faculty with librarian participation. For example, the campus student activism project led by a team of librarians at Duke University ([Hartsell-Gundy, Lawton, & Rozeat, 2020](#)) was self-contained and self-initiated. We found no mentioning of faculty involvement. Among limited literature where librarians took the lead and faculty were involved, project efforts tend to be on making existing library materials more accessible, such as the cataloging and digitization projects at University of Maryland ([Corlett-Rivera, 2017](#)), rather than creating new knowledge to fill a collection gap.

The collaboration, under this crowd-translation project, was carried out at an international and multidisciplinary, covering librarian, humanist and IT specialist groups from North America, Asia and Europe. And the librarian group took the lead, and made great outreach efforts, reaching out to the community of Ming scholars around the world. Traditionally, librarians' connection with scholars is limited within their own institutions. And humanities scholars tend to work alone rather than collaboratively as indicated by [Stone \(1982\)](#), and to have their own tight-knit circles per [Nicholas \(2015\)](#).

To solicit scholar collaborators and participants, the librarian team made tremendous outreach efforts, mainly through speaking and presenting at scholarly conferences and sending email to listservs. There were primarily four considerations when choosing which conferences to attend within project budget. They are (1) how many Ming scholars are likely to be there, (2) how likely they are capable of making (English translation) contributions, (3) whether conferences locations are diversified, and (4) whether conference dates work well with our project timeline. Finally, we decided to go to the following six conferences. The project outreach coordinator and leader took turn to attend and present there.

- The Forum on Forbidden City Research and Imperial Court's History of Ming and Qing Dynasties (August 2016, China)
- The 21st biennial conference of the European Association for Chinese Studies (August 2016, Russia)
- Annual meeting of the Society of Ming Studies (March 2017, Canada, March 2018, U.S.A.)
- Annual Meeting of the Society of Japan Historical Science (May 2017, Japan)
- An international conference on the Ming Dynasty and its Historical Era (September 2017, China)
- The 15th Biennial Conference of the Chinese Studies Association of Australia (July 2017, Australia)

In addition to these scholarly conferences, we seized every opportunity to present at professional conferences in the U.S., Canada, China and Korea, seeking help from librarian colleagues in spreading the words to their faculty. We also prepared project flyers in four languages (English, Chinese, Japanese and Korean) and distributed them at those conferences, asking to pass on the information in their network. Upon return, we also sent follow-up emails to associated listservs.

These outreach efforts served three purposes- networking, contribution solicitation, and (dictionary) usage promotion. Overall, our

project was well-received at all these conferences. Scholars agreed that the dictionary would be very useful to their research. Many of them showed interests in making contributions and/or using the dictionary upon its publishing. Although we do not have a tracking record showing to what extent these outreach efforts are correlated with final contributions, we do have registered scholars from North America, Asian and Europe in the crowd-translation system. And the bilingual dictionary has been linked from scholarly and professional websites from many parts of the world. Furthermore, through a networking activity during the early stage of the project at a conference in China, our outreach coordinator got to know the deputy director of the Chinese Society on Ming Dynasty History, who specializes in the government system. The senior historian became one of our expert consultants and made invaluable contributions to the project.

Employing crowdsourcing as a virtual collaboration method

We faced a great challenge in the translation part of the project. We had approximately 3500 government official titles in the beginning. About one fourth of them lacked English translations because many of them are complex and unusual. Those complex and unusual official titles are comprehensible only to highly specialized experts. However, translating a large quantity of various historical titles within a limited period is beyond the capacity of individual scholars who tend to have their deep silos of expertise and their own comfort zone when it comes to subject areas and time constraints. Providing English translations to government official titles from early Modern China also requires high-level language proficiency in both archaic Chinese and modern English. Furthermore, the community of Ming scholars is comparatively small, and has its own networking and communication circles. And its members are from all parts of the world, and may or may not be affiliated with Ming studies societies/associations in different countries and regions. As such, it is impossible for us to identify and solicit contributions directly from certain members or groups.

Technology offered a solution to collect widely distributed wisdom. After careful exploring literature and available options, we decided that a community-based knowledge creation by Ming scholars on an international scale, taking advantage of Web 2.0 technology, should be the route to take. It is only through crowdsourcing, that we would be able to maximize the number of qualified participants.

Crowdsourcing, since it was first coined and defined by [Howe \(2006\)](#), has been used for various activities and in different ways. Its definition is constantly evolving. Among numerous definitions is the more general and widely cited statement by [Estelles-Arolas and Gonzales-Ladron-de-Guevara \(2012\)](#).

“Crowdsourcing is a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit...”

Crowdsourcing is normally deployed to obtain information from a large amount of people on the Internet. However, literature also reveals that the crowd may come from the public as well as a community. Compared with the size of a community, the quality and qualification of community members, who would to respond and participate on a voluntary basis, is more important, just as [Modaresnezhad et al.](#) pointed out in their recent study ([Modaresnezhad, Lyer, Palvia, and Taras, 2020](#)).

“The attraction of the right crowd and their sustainable contribution are the keys to crowdsourcing success and require an understanding of the characteristics of the individual members of the crowd.”

Research has found that quality control is directly associated with the outcome of a crowdsourcing project ([Zhao & Zhu, 2014](#)). Crowd credential tends to be more important to the outcome of crowd-translation, because translation from one language to another is a complex cognitive task requiring not only proficiency in both languages but also cultural and/or domain expertise. Among a considerable amount of scholarship in the field of crowd-translation studies, Miguel Jiménez-Crespo’s work could serve as a starting point for overview of practices and issues in developing crowdsourcing translation projects. After a comprehensive analysis of existing models, industry and scholarly publications, the author proposes best practices for crowdsourcing translation quality, emphasizing “pre selection with exam”, “selection of the participants in the crowd”, “continuing evaluation of participant performance”, and so on ([Jiménez-Crespo, 2017](#)). For our project, the Ming studies community was chosen as the core crowd to meet academic standard. Members from this community worked together on the crowd-translation platform, to volunteer their knowledge for a common good that would benefit generations of Ming scholars, as well as members of larger research communities.

In addition to the right crowd, choosing or developing an effective crowdsourcing platform, including proper mechanisms for quality control, is directly linked to project success per [Ridge \(2014\)](#) and [Daniel, Kucherbaev, Capiello, Benatallah, and Allahbakhsh \(2018\)](#). Chris Callison-Burch’s research of assessing Amazon Mechanical Turk (MTurk) ([Callison-Burch, 2009](#)) also confirms the effectiveness of the quality control method. In this project, we surveyed a pool of open source software out there for crowdsourcing translation, such as ByBossa and Hive. However, none of them met our standards in terms of hierarchical organization of government agencies for official title browsing, anonymous peer-review and tiered administrative accounts. Ultimately, we determined to build our own for this project, and for other similar projects in the future.

The crowd-translation platform, which we have developed, has built-in quality control mechanisms, including registration and authentication for qualification, anonymous peer-review, and final selection by expert consultants. On the platform, we were able to collect English translations of thousands of Chinese historical official titles, many of which are unusual and obsolete, within just a year. This is an effective (high quality for complex texts) and efficient (short turnaround) accomplishment, providing an empirical evidence to the merit of crowdsourcing. Meanwhile, the empirical evidence has indicated that this crowd-translation system is a reliable digital tool for compiling bilingual dictionaries of any topic that requires in-depth domain knowledge and collective input.

Experiencing emerging roles of academic librarians

A literature review of the library science field three types of librarians’ new roles in the digital age that are relevant to our project. The first is building technical infrastructure for DH research. [Tzoc \(2016\)](#) presents four faculty-driven DH projects that were supported by Miami University Libraries’ Center for Digital Scholarship, where librarians played various roles, such as website design, usability test, user training, and more. [Arlitsch, Tzoc, and Millard \(2017\)](#) report their NEH grant project of building *Digital Scholarship Applications Dashboard*, a one-stop shop allowing faculty and students to easily choose, configure, deploy and evaluate common web applications for digital scholarship research. Whereas these technical infrastructures tend to be developed for faculty and students within their own institutions, the crowdsourcing system under the current project is for facilitating collaborative DH research without institutional boundaries. And the API (Application Program Interface) we’ve developed under Phase II provides a sustainable solution to the crowd-translation project.

The second is preparing and providing access to digital humanities texts on a given subject/topic, which is indisputably one of academic librarians’ strengths, as evidenced by the Rosarium Project ([Tryon,](#)

2017), the campus student activism project at Duke, as well as this current project. Our project started with around 3500 official titles collected by the CBDB team through text-mining of digital biographical resources. We identified and added about 2000 more, from significant Ming historical documents, in either digital or print form as listed at <https://mingofficialtitles.lib.uci.edu/#/bibliography>. We also spent a great amount of time on verifying and cleaning the computer-extracted data from CBDB. To make the final set of 4552 official titles better organized and more usable for online access, we applied metadata management expertise to develop a hierarchical office structure of three-tiers to pair with the official titles, building on historical research, and feedback from our expert consultants.

The third new role is about transforming project management skills. Both Currier, Mirza, and Downing (2017) and Brandenburg et al. (2017) argue that librarians can leverage their project management skills from different functional units of a library to form a holistic mindset for managing DH projects. Under this project, in addition to utilizing technology (i.e., the crowdsourcing system) to manage the international collaboration on translating Ming official titles, several library units at UC Irvine were involved in managing personnel, financial, and technical aspects of the project. The Human Resource unit helped to determine the job titles and salary ranks for the contract programmer and research specialist, and to bring the two external hires on board. The Business Office prepared and shared with the project leader, on a regular basis, grant reconciliations. The IT unit provided strong technical support on arranging server space and assisting in system development. Various documents, such as timeline, decision tracker, task list and workflow, were created and kept on library wiki to help the project proceed as planned, on schedule and within budget.

Actually, librarian's strengths in digital resources, user-computer interaction, information organization and representation, and project management make us ideal DH research collaborators and coordinators, as implied by the DH Stacks model by Berry and Fagerjord (2017). Librarians may partner with scholars in various types of DH research, with regards to digital data, metadata, tool use, copyright, publication, and many more.

Apart from contributing our skills in technical infrastructure construction and project management, our experience with the dictionary project demonstrated that academic librarians could be involved in DH scholarship in depth, through taking more advanced roles. Examples of advanced roles include content creator and publisher, research partner, project leader, hybrid scholar, and grant writer. Through working closely with Ming scholars and expert consultants, and through learning about Ming political system from related historical literature and research work, we experimented with expanding our roles from collection curation to knowledge creation, having compiled and published the *Chinese-English Dictionary of Ming Government Official Titles*. Meanwhile, we went beyond the research partner role, to lead the DH project, to coordinate collaboration among international scholars, and to oversee project execution.

External grants are essential to the success of this project. To be a successful grant writer, one needs to have these qualities: a creative idea, a promising value, a rigorous and feasible method, a qualified and dedicated researcher (or research group), a commitment to the proposed project, as well as a good match between project objective and mission of a grant agency. Our team is fortunate that we have submitted our grant applications to the right funding agencies- Andrew H. Mellon Foundation, who "believes that the arts and humanities are where we express our complex humanity" (Andrew W. Mellon Foundation, 2021); and Geiss Hsu Foundation, whose mission is to "encourage and sponsor scholarly research and interpretation of the Ming Dynasty (1368-1644) in China" (Geiss Hsu Foundation, n.d.).

Learning through doing

While leading this complex project, we faced many challenges,

including the lack of advanced subject knowledge, difficulty of cultivating trust and contribution from scholar, constraints of technological and financial resources, competing work and research schedule, and conflicting perspectives among team members. Over the one-year project period, these challenges turned out to be opportunities for us to learn and grow. We learned how to compromise between project outcomes and technology/budget constraints, to balance time spent on our library work and this DH project, to build consensus over various issues ranging from project boundary, structure, categorization and organization of official titles, to specific official names. We succeeded in overcoming these challenges and bringing our project to light, thanks to (1) incredible support from Ming scholars, especially the expert consultants, (2) generous funding from grant agencies, (3) helpful assistance from UC Irvine Library administrative and IT teams, (4) hard and solid work of computer programmer and research specialist, as well as (5) trust, respect and collaboration among members of different teams.

Meanwhile, as academic librarians, who do not have intensive knowledge about a historical government system from hundreds of years ago, we acquired essential and specialized knowledge through reading academic publications and learning from scholar collaborators. Likewise, by taking the learning-by-doing approach, we came closer to an understanding of digital scholarship, and gained valuable experience in DH project management and team collaboration. We also built skills in compiling and publishing a subject-specific dictionary. All these activities and accomplishments helped strengthen our confidence in diving deeper into digital scholarship.

Conclusion

This librarian-led collaborative project has produced not only the crowdsourcing system and the authority file respectively as the technical and information infrastructure for supporting DH research, but also the bilingual dictionary to fill a publication gap for Chinese history scholars and students. Although the project has officially ended, we are committed to an ongoing enrichment and enhancement of the dictionary through a technology-empowered sustainable solution for ongoing text-mining and crowd-translating additional Ming official titles.

As this project demonstrates, there are a variety of new roles that college and research librarians could explore and undertake in the realm of digital humanities and throughout the lifecycle of scholarly communication. And only through constantly exploring and undertaking new roles, can we sustain and thrive in the ever-evolving digital scholarship.

To dive deeper in DH research, academic librarians first need to be proactive and motivated, being willing and ready to take advanced roles. Through this current DH project, we succeeded in expanding our role from curators of collection to creators of knowledge. Traditionally, we build our collection based upon existing publications. As to a patron request for a Chinese-English dictionary of Ming government official titles, we might have no choice but to say ten years ago, "Sorry, there is no such a dictionary!" But nowadays, taking advantage of advanced technologies, our answer could be more proactive and promising, "Well, there is none; but let's build one together!"

To reposition librarians' role from research supporter to research partner, we need to stay connected with scholars and be needs-oriented. Compared with the assumption of "if we build it, they will come", the "we build for your needs" approach is more effective, because scholars are more willing to collaborate if what we do is perceived as useful to them.

As DH research is collaborative in nature and normally involves various teams of different background, it is critical that all team members are open-minded, communicative and respectful. Debates and arguments over different ideas and approaches may not be a bad thing. Instead, with all-inclusive open-mind and respect, debates and arguments may contribute to a more closely-examined outcome. Also important is to have a project leader, who is fair and decisive, and domain experts, who are knowledgeable and accountable.

Furthermore, digital humanities research cannot be done without funding. External grants are essential to busy librarians who want to explore and grow in new areas, which require tremendous time and subject/technology expertise. We are so thankful to Andrew W. Mellon Foundation and Geiss Hsu Foundation for their support. It is hard to imagine that we could have led the DH project to a success, without their generous support.

Last but not least, publishing with open access (OA) leads to increased dissemination and impact of research. Academic libraries and librarians are in the center of the OA movement, pushing forward its agenda. Through making our own research products, including the bilingual dictionary and the crowd-translation system, available for everyone, we serve not only as an advocate and a venue provider, but also a content contributor.

CRedit authorship contribution statement

Conception and design of study: Y. Zhang

Acquisition of data: Y. Zhang

Analysis and/or interpretation of data: Y. Zhang, S. Xue, Z.H. Xue

Drafting the manuscript: Y. Zhang, S. Xue, Z.H. Xue

Revising the manuscript critically for important intellectual content:

Y. Zhang, S. Xue, Z.H. Xue

Approval of the version of the manuscript to be published: Y. Zhang, S. Xue, Z.H. Xue

Acknowledgments

The authors would like to thank all Ming scholars, expert consultants, IT specialists and our research specialist Dr. Li Ni for their directly contributions to the project. We thank Harvard's CBDB (China Biographical Database) team and UC Irvine Libraries' administration and IT unit for their support. We thank Xiaohe Ma of Harvard Yenching Library for his contribution to the project, during his participation as a project team member in the early stage. We would also like to thank the Andrew W. Mellon Foundation and the Geiss Hsu Foundation for funding support.

References

- Anastasiou, D., & Gupta, R. (2011). Comparison of crowdsourcing translation with machine translation. *Journal of Information Science*, 37, 637–659.
- Andrew W. Mellon Foundation. (2021). Mission. Available online at <https://mellon.org/about/mission/>.
- Arlitsch, K., Tzoc, E., & Millard, J. (Jul, 2017). An on-demand and cloud-based digital scholarship applications dashboard. *Journal of Library Administration*, 57(5), 563–576.
- Association of College & Research Libraries. Principles and strategies for the reform of scholarly communication. Last modified June 24, 2003. Available online at <http://www.ala.org/acrl/publications/whitepapers/principlesstrategies>.
- Association of College & Research Libraries. Scholarly communication toolkits: Scholarly communication overview. Available online at <https://acrl.libguides.com/scholarlycommunication/toolkit>.
- Berry, D. M., & Fagerjord, A. (2017). *Digital humanities: Knowledge and critique in a digital age*. Polity Books.
- Blanke, Tobias, Kristel, Conny, & Romary, Laurent (2017). Crowds for clouds: Recent trends in humanities research infrastructures. In Agiati Benardou, Erik Champion, Costis Dallas, & Lorna Hughes (Eds.), *Cultural heritage infrastructures in digital humanities* (p. 60). Routledge.
- Brandenburg, M. D., Sigrid, A. C., Joque, J., MacEachern, M. P., & Song, J. (2017). Interdisciplinary collaboration: Librarian involvement in grant projects. *College and Research Libraries*, 78(3). Available online at <https://doi.org/10.5860/crl.78.3.272>.
- Callison-Burch, C. (2009). Fast, cheap, and creative: Evaluating translation quality using Amazon's Mechanical Turk. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing* (pp. 286–295).
- CBDB. China Biographical Database Project (CBDB). Available online at <https://projects.iq.harvard.edu/cbdb/home>.
- CBDB. Contributors. Available online at <https://projects.iq.harvard.edu/cbdb/collaborators>.
- Corlett-Rivera, K. (Apr, 2017). Subject librarian as coauthor: A case study with recommendations. *College & Undergraduate Libraries*, 24(2–4), 189–202.
- Council on East Asian Libraries. (2016). Mellon grant recipients for 2016, June 15. Available online at <http://cealnews.blogspot.com/2016/06/mellon-grant-recipients-for-2016.html>.
- Currier, B. D., Mirza, R., & Downing, J. (Apr, 2017). They think all of this is new: Leveraging librarians' project management skills for the digital humanities. *College & Undergraduate Libraries*, 24(2–4), 270–289.
- Daniel, F., Kucherbaev, P., Cappiello, C., Benatallah, B., & Allahbakhsh, M. (2018). Quality control in crowdsourcing: A survey of quality attributes, assessment techniques, and assurance actions. *ACM Computing Surveys*, 51(1), 1–40.
- Estelles-Arolas, E., & Gonzales-Ladron-de-Guevara, F. (2012). Towards an integrated crowdsourcing definition. *Journal of Information Science*, 38(2), 189–200.
- Geiss Hsu Foundation. Mission. Available online at <https://geissfoundation.us/about/mission-statement/>.
- Golomb, L., Braunstein, L., & Hartsell-Gundy, A. (2015). *Digital humanities in the library: Challenges and opportunities for subject specialists*. Chicago: Association of College and Research Libraries.
- Hartsell-Gundy, A., Lawton, K., & Rozear, H. (Jan, 2020). 17 librarians and one big undertaking: Creating a digital project from start to finish. *Journal of Electronic Resources Librarianship*, 32(1), 19–28.
- Howe, J. (2006). The rise of crowdsourcing. *Wired*, 14(6), 1–4.
- Hucker, C. (1985). *A dictionary of official titles in Imperial China*. Stanford University Press.
- Jiménez-Crespo, M. A. (2017). *Crowdsourcing and online collaborative translations: Expanding the limits of translation studies* (p. 133). John Benjamins Publishing Company.
- Kear, R., Joranson, K., & Humanities, D. (2018). *Libraries, and partnerships: A critical examination of labor, networks, and community*. Cambridge: Chandos Publishing.
- Kennedy, G. A. (1983). The relationship between acquisitions and collection development. *Library Acquisitions: Practice and Theory*, 7, 225–232.
- Kirschenbaum, M. G. (2010). *What is digital humanities and what's it doing in English departments*. 150 p. 56). ADE Bulletin. Available online at: www.ade.org/bulletin/.
- Li, B. Z. (2021). *An early modern economy in China the Yangzi Delta in the 1820s*. Cambridge, U.K.: Cambridge University Press (p. xvi).
- Modaresnezhad, M., Lyer, L., Palvia, P., & Taras, V. (2020). Information Technology (IT) enabled crowdsourcing: A conceptual framework. *Information Processing and Management*, 57(2), Article 102135.
- Nicholas, D. (2015). Using, citing and publishing scholarly content in the digital age: Case study of humanities researchers. *ZIN (Zagadnienia Informacji Naukowej. Studia Informacyjne)*, 53(1), 7–19.
- Nielsen, J. (2006). *The 90-9-1 rule for participation inequality in social media and online communities*. Available online at <https://www.nngroup.com/articles/participation-inequality/>.
- Papelitzky, E. (2019). *Editing, Circulating, and Reading Huang Zhong's Hai yu 海語*. *Ming Qing Yanjiu*, 23, no.1 (pp. 1–38). Brill.
- Qian, L. X. (2020). Distribution maps of Chinese poets in the Ming dynasty (1368–1644): A geographical visualization experiment. *Library Trends*, 69(1), 289–307.
- Ridge, M. (2014). *Crowdsourcing our cultural heritage*. Farnham, Surrey, England: Ashgate.
- Stone, S. (1982). Humanities scholars: Information needs and uses. *Journal of Documentation*, 38(4), 292–313.
- The Geiss Hsu Foundation. (2019). Past awards. Available online at <https://geissfoundation.us/past-awards/?pg=2>.
- The Ming History English Translation Project. More resources on great Ming. Available online at <https://kmit.ucsd.edu/minghistoryinenglish/other-ming-resources/>.
- Tryon, J. R. (Apr, 2017). The Rosarium project: A case of merging traditional reference librarian skills with digital humanities technology. *College & Undergraduate Libraries*, 24(2–4), 171–188.
- Tzoc, E. (2016). Libraries and faculty collaboration: Four digital scholarship examples. *Journal of Web Librarianship*, 10(2), 124–136.
- University of British Columbia. Documentation: Guide to Ming studies. Available online at https://wiki.ubc.ca/Documentation:Guide_to_Ming_Studies.
- Zhang, Y., Liu, S., & Matthews, E. (2015). Convergence of digital humanities and digital libraries. *Library Management*, 36(4/5), 362–367.
- Zhao, Y. X., & Zhu, Q. H. (2014). Evaluation on crowdsourcing research: Current status and future direction. *Information Systems Frontiers*, 16, 417–434.