10-3-2014

A Crowdsourced Library Help System

Tao Zhang
Purdue University, zhan1022@purdue.edu

Ilana Stonebraker
Purdue University, stonebraker@purdue.edu

Recommended Citation

Final Report for IMLS Grant Number LG-46-13-0239-13

1. Administrative Information
Institution: Purdue University
Project Title: A Crowdsourced Library Help System
Award Amount: $23,831.00
Total Project Cost: $29,072.65
Grant Start Date – Grant End Date: 07/01/2013 – 06/30/2014
Project Director Name: Ilana Stonebraker (previously Ilana Barnes), Tao Zhang

2. Project Summary
Help information regarding library resources and tools is a critical part of library services, but existing help channels are often fragmented and under-maintained. Library help needs to have a centralized and engaging platform for users to receive assistance and experts including librarians, faculty, and students to share knowledge and experience. To meet this critical need, we have developed CrowdAsk, an open-source help system for academic libraries based on the concept of crowdsourcing. CrowdAsk supports voting of questions and answers by users as well as scores and badges for user motivation. Our implementation and evaluation with undergraduate classes at Purdue University showed that CrowdAsk is effective in meeting users' information needs beyond traditional library reference help. Users have asked various types of questions and reached high levels of scores and badges in one semester's time. We have integrated CrowdAsk with existing reference service and websites in Purdue University Libraries to continue developing the crowdsourcing community after the grant.

3. Process
A. Requirements: development, refinement and deployment (August to September 2013)
The first project activity was to develop detailed technical specifications and use cases for CrowdAsk. Zhang and Stonebraker wrote a 6 page long Functional Requirements document. This documents included information on mechanics that they believe should be implemented in CrowdAsk, including user roles, gaming mechanisms, and user interface specifications. To create this document, Zhang and Stonebraker relied heavily on their previous research on help system metrics[1], and by reviewing existing crowdsourced help systems such as Stack Overflow. Zheng then reviewed existing open source code for questions and answer systems and identified a strategy for system implementation based on the technical specifications and use cases from Stonebraker and Zhang. Question2Answer[2] was selected as an acceptable starting point for the project as it had many of the elements that had been highlighted in the Functional Requirements, with the exception of the gamifications and librarian user roles. By using this open source code

---

[1] Zhang, T., Barnes, I. R., & Promann, M. (2014). Building better help: User characteristics' effect on library help design. *Communication Design Quarterly*, 2(2), 21-28.
[2] Question2Answer source code: https://github.com/q2a/question2answer

as a base, the group could save time in project implementation and also allowed the project to build upon existing knowledge in the open source community.

B. CrowdAsk System beta: Development and Design (September to January)

From September to December, Zheng developed the database structure, procedures for data storage and retrieval, and a working interface for user testing. Meanwhile, Promann developed a design document, including colors and logos. During this period, the group met weekly to go over changes, give updates and feedback. Both graduate student assistants engaged in rapid prototyping and testing efforts throughout the project as part of the agile user-centered development methodology.

C. Weekly meetings (ongoing during project period)

Every week all staff met and reported progress, discussed design and implementation issues, and planned for future tasks. Meeting minutes were posted on the project's internal site. All project staff shared information and communicated regularly through email and a dedicated project site hosted by Purdue University Libraries.

D. Beta Testing with English 106, MGMT 175 and GS 175 (January to April 2014)

During September to December, Stonebraker recruited three English 106 sections, four MGMT 175 sections (including one she taught) and one GS 175 section for implementing CrowdAsk in the Spring 2014 semester. In order to spur engagement, students were awarded a small amount of extra credit for reaching certain point levels on CrowdAsk. To answer questions and introduce CrowdAsk, Stonebraker visited six class sessions at the begining of the semester, then returned to those sessions six weeks later to answer any questions that had arose. During this period, instructors and students reported bugs that were addressed immediately or discussed in weekly meetings. Since CrowdAsk was intuitive to use, training of instructors was not needed.

E. Usability testing of CrowdAsk prototype (February to March 2014)

Due to scheduling issues, usability testing of CrowdAsk started in March, versus January as outlined in the project plan. In February, IRB was sought for usability testing of CrowdAsk. In March, four user tests were conducted. These users were recruited both from the beta-test classes and outside of the best-users group. Students in the beta-test group who were particularly active on CrowdAsk were recruited in order to further understand why they participated beyond the levels needed for extra credit. Results from these usability tests were reported back in the weekly meetings, and potential usability issues identified from the tests were addressed.

F. Release of CrowdAsk source on GitHub (July 2014)

The project plan listed that the CrowdAsk code would be available open source for others to download and adapt. By making CrowdAsk open source, other libraries could potentially adapt our system and integrate it with their web presence, thus creating a broader impact for enhancing

existing library reference services. The code was release on July 1, then updated on July 23 and is available on GitHub[3].

G. Publications, press and presentations related to CrowdAsk (July to November 2014)
As mentioned in the project proposal, Stonebraker and Zhang disseminated project results in both user experience and reference librarian's publications. CrowdAsk was presented internally twice at the Purdue University Libraries Learning Council. In addition, it was presented nationally as a lightning talk at the 2014 ALA Annual Conference in Las Vegas at the LITA Heads of Technology Interest Group as well as presented as part of Stonebraker's presentation on flipped learning and business information literacy at the BRASS Business Reference in Academic Libraries Forum. It was also presented at the Indiana University Information Literacy Colloquium in August. Stonebraker and Zhang have contributed a book chapter on CrowdAsk to "Reimagining Reference in the 21st Century", which will be published in conjunction with the 2014 Charleston Conference in November. Stonebraker will also speak at the Charleston Conference as a plenary presentation on crowdsourcing library systems. Stonebraker and Zhang have been accepted to speak at Internet Librarian 2014 in October. A contributed paper on CrowdAsk has also been accepted to ACRL 2015.

4. Project Results
A. What we have completed
As described in our project proposal, the expected outcomes of this project include: (1) a functional library help system (CrowdAsk) integrated with Purdue University Libraries website; (2) source code of CrowdAsk released to the public; (3) user evaluation results; (4) user surveys and web analytics data; and (5) scholarly publications and conference presentations related to CrowdAsk.

CrowdAsk has been implemented on Purdue University Libraries website on the Archives and Special Collections (ASC) homepage[4] since July 2014 and on the Ask a Librarian page[5] since August 2014. The ASC operates a main reference account (spcoll@purdue.edu) that receives nearly 40 new inquiries per month (2013-2014 statistics). The most-asked questions include queries about content use and rights in publications, queries regarding well-known collections, and scholarly student questions. Via CrowdAsk, ASC hope to address frequently the frequently asked questions and provide a moderated community for all users. CrowdAsk has a category for ASC-related questions that is linked to the feature box on ASC homepage. As of August 2014, there are 42 questions in the ASC category, with 40 questions with one answer each. The Ask a Librarian page is the main entry point for the digital reference services of Purdue University Libraries. CrowdAsk has been marketed as a secondary choice for users to get help, especially when the online chat is not available. The Purdue University Libraries reference staff have

---

[3] CrowdAsk source code: https://github.com/crowdask0/crowdask
[4] Archives and Special Collections of Purdue University Libraries: https://www.lib.purdue.edu/spcol
[5] Ask a Librarian of Purdue University Libraries: https://www.lib.purdue.edu/help/askalib

agreed to provide limited support for CrowdAsk to make sure questions are answered in a reasonable time frame. The staff have also agreed to evaluate the time and resources used to support CrowdAsk and provide periodic feedback to the CrowdAsk projec team. The Purdue University Libraries web team has agreed to monitor for potential technical issues of CrowdAsk. When a technical issue is identified, the web team will provide assessment of time and resources needed to address the issue. The CrowdAsk team will then work with the web team to develop feasible solutions in accordance with the web team's development plan. As of August 25, 2014, there are 14 questions posted by general users on CrowdAsk. With the support from ASC, digital reference, and the libraries web team, CrowdAsk has a very good starting point to develop sustainable user engagement and community involvment as part of the Purdue University Libraries website.

The source code of CrowdAsk created by the project is available under GNU General Public License version 2 on GitHub[3]. As stated, CrowdAsk is based on the open source Question2Answer system[2] with further development and customization. Table 1 provides a brief summary of CrowdAsk's features. The features are integrated in an installation script for easy implementation at other institutions. We have also prepared a cookbook as part of the source code for administrators to set up and manage the system. The easy availability of the code and issue tracker will facilitate the contribution of third parties to the ongoing maintenance and enhancement of CrowdAsk.

Table 1. Summary of CrowdAsk's features.

User roles and points	Each user role is associated with a set of privileges. An anonymous user can only browse and view questions, answers, and comments. By default, all registered users have the following privileges: 1) ask new questions; 2) answer existing questions; 3) post comments to questions and answers for clarifications; 4) mark an answer to his/her own question as accepted; 5) vote up and down answers; 6) create tags and select categories; 7) set bounties for questions to encourage answers. Whenever a user asks a question, answers a question, or votes on a question or answer, he receives points. He also receives points for his answer being selected as best. The point values can be adjusted on the administrative interface of CrowdAsk. The system will automatically assign additional privileges to a user based on the user's points, "leveling up" that user to a new user role. Additional privileges include closing votes of the user's question, reopening votes of a question, and locking any question so the question will not accept new answers. A moderator (designated by the administrator of CrowdAsk) has the ability to close and reopen votes of a question or answer. Note that if the vote for a question is closed, the votes of the question's answers are also closed.
Questions and answers	A question can have multiple answers, but only one answer can be marked as accepted (best answer). An authenticated user can mark a question as her

	favorite. A question can be locked to prevent unconstructive and off-topic discussions. An administrator can also hide questions from public view. A question can be protected so that the question will not accept votes. The list of answers is ordered by up votes by default and also can be ordered by time the answers were posted or edited, number of views, and number of answers.
Tags and categories	An authenticated user can assign existing tags to the question she asks or create her own tags organically. Categories support a hierarchical structure for questions and user roles. Usually tags are subjects and categories that reflect communities or groups of users, such as a class or department. An authenticated user can assign or modify an existing category to the question the user created. A question has none or only one category. A category is associated with a list of questions in the category. Users can use categories to browse related questions. The moderator role can be associated with one or more categories.
Badges	In order to encourage positive behavior, badges also are included in CrowdAsk. In the administrative interface, administrators can create an unlimited number of badges. We have implemented 23 badges in our system. Some example badges we have implemented include "Knowledgeable" (having at least one answer selected as best answer); "Commentator" (left at least 10 comments); "Good Answer" (answer with more than 25 up votes); and "Outstanding Contributor" (left 25 or more comments, and answered 20 questions or more).
Bounties	A user can assign a bounty to a question when the user first creates the question. A user must have at least 120 points to assign a bounty. Bounty points for a question can be set from 20 to 100. Once the user has set a bounty, the system deducts the point from the user and stores the point in a secured table. All bounties are paid for upfront and are nonrefundable under any circumstances to prevent students from revoking their bounties after getting the answer they wanted, and thus gaming the system. Once the user selects a best answer for the bounty question, the answerer gets the bounty points.

B. Implementation results

In the spring 2014 semester (January 13 to May 18), we implemented CrowdAsk with three undergraduate courses at Purdue, including English 106 (First-Year Composition), Management 175 (Information Strategies for Management), and General Studies 175 (Information Strategies for Hospitality and Tourism Management). English 106 is a required course for all incoming students, and it is aimed at providing students essential skills for researching a topic and presenting the topic in various written formats. Management 175 and General Studies 175 are required courses designed to develop students' information-gathering, evaluation, and presentation skills in the two disciplines. In total these three courses included 12 class sessions and over 370 students. We introduced CrowdAsk to the students at the beginning of semester and frequently encouraged students to use CrowdAsk in the classes when they needed help researching a topic. We added the link to CrowdAsk in the course management system as part of the course resources. The overall usage data of CrowdAsk between January 5, 2014, and April 2, 2014, is provided in Table 2.

Table 2. Usage data of CrowdAsk, January 5, 2014, to April 2, 2014.

Item	Amount
Registered users	211
Users who posted (questions or answers)	99
Users who voted	106
Questions	122
Answers	232
Question owners	59
Question answerers	72
Comments	14
Most votes	35
Most answers	8
Most views	92
Unanswered	5
Highest user score (registered users start from 100)	2,080
Highest number of badges	11

Data from Google Analytics showed that from January 5, 2014, to April 2, 2014, CrowdAsk had 1,150 visits from 474 unique visitors. The total number of page views was 14,715, with average 12.8 pages per visit. The average visit duration was 6 minutes and 7 seconds. This shows that CrowdAsk attracted a good amount of traffic from the classes. The students were engaged with CrowdAsk as they stayed and viewed some amount of pages for an average visit.

Preliminary analysis of the question titles and answer texts showed that users asked the following categories of questions on CrowdAsk, ordered by their frequency (Table 3).

Table 3. Categories of questions on CrowdAsk.

Question Category	Definition	Example Questions
Course-related	Questions that ask for information about particular	"Do you know how to retake the

	courses.	quiz on blackboard (for Management 175)?" "How much will my writing improve throughout the semester?"
CrowdAsk-related	Questions that are about the point system, badges, and type of questions on CrowdAsk.	"How do you earn points on CrowdAsk?" "May I know the full list of badges and how to achieve them?" "Are we only allowed to ask academic-related questions here?"
Library services or resources	Questions that are about where to find certain information and how to access physical and electronic resources.	"Is there a way to search the libraries catalog just for movies?" "How do I reserve a study room at library?" "How do you get the actual article to come up on Business Source Premier instead of just the abstract?"
How-to	Questions that ask for instructions.	"What is a good website to use to do a voiceover on Prezi?" "How to analyze the financial tables of a company?"
Conceptual	Questions that are conceptual, abstract, and do not involve specific contexts.	"What is the best citation management software?" "Could someone tell me what is the meaning of APA citation?"

Most questions on CrowdAsk fit into one of these categories, but some questions could be assigned to more than one category, for example, "How do I find the DOI for an IBISWorld article?" This question was asked by a student in the Management 175 class, but the answer could be useful for other library users.

C. User test results

We recruited four students and three faculty members for the user evaluation of CrowdAsk. Three students were from the current users and received high points in the system; the other student and the three faculty members did not use CrowdAsk before the test. The test tasks included: 1) registration and login; 2) ask a question and assign tags to the question; 3) propose an answer to an existing question; 4) select the best answer to the question asked earlier; 5) view user profile and understand user privileges; and 6) explore other features of CrowdAsk, including tags, categories, and site search. Participants explored CrowdAsk before they performed the six test tasks. During the test tasks, they were encouraged to talk aloud about their expectations, difficulties, and general comments about using CrowdAsk. After the tasks, participants responded to open questions about their experience using CrowdAsk.

Overall, participants' task performance showed that basic tasks—such as browsing existing questions, asking and answering a question, selecting the best answer, and voting for questions and answers—were easy and without major obstacles. Participants completed the test tasks within reasonable time duration (averaging approximately two minutes) and did not require assistance. We observed that participants tended to not assign tags to their questions. The major reason was that they did not use tags in general and did not find tags particularly useful for them. This suggests that proper intervention (e.g., adding/editing tags) may be necessary during the initial launch of CrowdAsk. Participants commented that the meaning of badges and user levels were not clear to them. Badges are shown on user profile pages, but the user needs to move the mouse cursor over each badge in order to see its corresponding rule. Similarly, user level and privilege information is shown on the user profile page, but not on the list of questions page. Two experienced users of CrowdAsk did not know that they have additional privileges (e.g., editing questions and answers) due to their high points in the system. Consequently, we noticed that only a few users used their editorial privileges on CrowdAsk. User participation in the content management will be critical if CrowdAsk continues to grow as a user community. Therefore, we have proposed new interface designs to address these issues and promote awareness of features such as points, badges, and privileges among CrowdAsk users.

We also collected user feedback on CrowdAsk through course evaluations. In general, students liked the idea of asking questions and helping each other on CrowdAsk, without emailing teachers all the time. A few students thought they did not have questions to ask, and some questions and answers on CrowdAsk were vague, which reinforces the importance of user guidance and editorial intervention (i.e., clarifying questions and answers through editing and comments, as well as building useful content).

The above usage data and evaluation findings indicate that CrowdAsk could be an effective tool to meet users' information needs beyond traditional library reference help. Users have asked various types of questions and reached high levels of scores and badges in a relatively short time period. With the integration of CrowdAsk on Purdue University Libraries website, we will

continue to develop useful content, engage users, and assess CrowdAsk through content analysis and user testing.

D. What we have learned

From the project, we learned that involving the "crowd" is critical for CrowdAsk to grow and be successful. Introducing CrowdAsk to classes and providing extra credits for active students in the classes have been proved to be an effective way to build the initial user engagement. However, outside the classes, it will still be a challenge to attract general library users, particularly expert users, to consume information and contribute back to CrowdAsk. On the other hand, the majority of questions the digital reference staff encounter are about item location, article access, directions, and library policies. Those questions usually take a relatively short time to address and we expect questions on CrowdAsk to be similar. The minimal effort needed to vote or answer a simple question on CrowdAsk should not be an obstacle for the majority of users. For difficult research questions or questions about a particular subject, the digital reference staff could provide answers or share the questions with other subject experts.

In addition, we learned in usability testing users described being more motivated by reciprocity and community building than by gamification elements. When asked why they contributed to CrowdAsk, they described their contribution in terms "helping each other out" especially helping those that helped them. The points and badges, while interesting to them, was not as motivational to students as much as extra credit or reciprocity with other classmates and colleagues.

We also learned that in CrowdAsk's academic context, it is important to show the authority of answers from librarians and faculty members. The faculty members in our user evaluation were concerned that students may trust potentially incomplete or incorrect answers on CrowdAsk in their research. While we expect the voting mechanism and user points could mediate the trust issue, we have proposed interface improvement to explicitly show the roles of librarians and faculty members on CrowdAsk. In addition, the digital reference staff have been trained to educate and support students' research, not simply providing direct answers to specific questions. Users can also flag inappropriate content on CrowdAsk to the moderators and administrators. With these measures in place, we will closely monitor the content of questions and answers on CrowdAsk and provide timely intervention when necessary.

E. What's next?

The IMLS Sparks! Grant has enabled us to build a strong foundation of CrowdAsk and understand user interactions. As mentioned, Stonebraker and Zhang will work closely with the virtual reference staff and web team in Purdue University Libraries to further support and assess CrowdAsk. Stonebraker will continue to implement CrowdAsk in selected Purdue classes. Zhang will set up and monitor metrics on Google Analytics to assess user engagement. Additional user tests will be scheduled from the general users of Purdue University Libraries and ASC websites. Through the open source code on GitHub and our dissemination efforts, we will actively seek opportunities to further enhance CrowdAsk and develop different implementation models.

F. Resources
CrowdAsk production site: https://sites.lib.purdue.edu/crowdask/
CrowdAsk source code and documentation: https://github.com/crowdask0/crowdask

Project-related publications
1. Stonebraker, I., & Zhang, T. (2014). CrowdAsk: Crowdsourcing Reference and Library Help. In David A. T. & John G. D. (eds.), *Reimaging Reference in the 21st Century*, Purdue University Press.
2. Stonebraker, I., & Zhang, T. (2014). CrowdAsk: A Crowdsourced Library Help System. Lightning Talk Presented at American Librarian Association: Library and Technology Association Heads of Technology Interest Group, Las Vegas, NV.
3. Stonebraker, I., & Dugan, M. (2014). Flipping Business Information Literacy. Talk Presented at American Librarian Association: Business Reference In Academic Libraries Forum, Las Vegas, NV.
4. Gibson, C., & Stonebraker, I. (2014). Improving Instruction: Metaliteracy Through Crowdsourcing in the Classroom. Talk Present at the Indiana University Libraries Information Literacy Colloquium, South Bend, IN.