

Tell Me About My Family: A Study of Cooperative Research on Ancestry.com

Heather Willever-Farr
The iSchool at Drexel
Philadelphia, PA 19104
215-895-2474

hlw29@drexel.edu

Lisl Zach
The iSchool at Drexel
Philadelphia, PA 19104
215-895-2474

lisl@drexel.edu

Andrea Forte
The iSchool at Drexel
Philadelphia, PA 19104
215-895-2474

aforte@drexel.edu

ABSTRACT

Q&A forums for the exchange of genealogical information are becoming increasingly common on the web. Yet, relatively little is known about the socio-technical dimensions of genealogists' interactions in such forums. This study examined exchanges between genealogists on a popular Q&A message board on Ancestry.com. Our findings suggest that the web context shapes the types of exchanges and cooperative activities in which genealogists engage. Research has found that in face-to-face exchanges genealogists tend to help other genealogists by providing instructional guidance both on a one-to-one and a many-to-one basis. Our findings suggest that the presence of online genealogical data and the affordances of interactive computer technologies may be pushing answerers away from providing instruction on how to find family history data and pushing them toward providing those data outright. Answerers worked cooperatively to provide family data, suggesting that the web context is leading many genealogists to engage in cooperative research not collaborative instruction.

Categories and Subject Descriptors

K.4.m. [Computers and Society]: Miscellaneous.

General Terms

Human Factors

Keywords

Q&A Websites, Social Reference, Social Informatics, Genealogy, Genealogists, Family History.

1. INTRODUCTION

People turn to the web not only to learn about others, but to learn about themselves. According to the PEW Internet and American Life Project, 54 million Americans belong to a family where someone in the family has used the Internet to research their family history or genealogy [13]. This is not surprising given the growing amounts of family history data that have been digitized, indexed, and made available on the web. To support the large numbers of genealogists who are online, web-based Q&A forums have cropped up for those who are seeking and sharing family

history information. While the use of genealogy Q&A forums is wide-spread, relatively little is known about the interactions among users of these forums.

Existing literature on genealogists' information behaviors has only touched upon this phenomenon, leaving a vague picture of genealogists' interactions on the web. In addition, we know of no existing study of Q&A websites and user behavior that explores the interactions of genealogists as they construct family histories. Recent work on Q&A websites and user behavior has focused on large multipurpose Q&A websites, but has seldom explored social interactions in specialized Q&A forums for well-defined communities with long-standing practices, such as genealogists. This gap in the literature raises the following questions:

- 1) What kinds of social interactions emerge among askers and answerers on a message board for genealogists?
- 2) In what ways do the mediated environment and existing practices of genealogists influence those interactions?
- 3) Are Q&A forums serving as a means to educate users about the practices of a specific community of practice, in this case, genealogists?

To answer these questions, we analyzed message posts on a heavily used Q&A message board on the popular genealogy website, Ancestry.com.

2. LITERATURE REVIEW

2.1 Social Aspects of Family History Construction

Relatively little is known about the socio-technical aspects of genealogists' interactions in web environments; however, existing research on genealogists' information behaviors, primarily in the context of physical archives and face-to-face encounters, may serve as a baseline for understanding how the web is influencing genealogists' information behaviors in Q&A forums.. Duff and Johnson [6] studied amateur genealogists' use of in-house archival finding tools and their face-to-face interactions with information professionals. They identified three different stages in the genealogical research process: 1) collecting names of family members; 2) gathering detailed information on family members; and 3) contextualizing the detailed information by learning about broader history. In addition, Duff and Johnson found that genealogists often worked around archival access systems and relied more on their own social networks than on information professionals to obtain information.

Copyright is held by the author/owner(s).

iConference 2012, February 7–10, 2012, Toronto, Canada.
ACM 978-1-4503-0782-6/12/02.

Building on Duff and Johnson's stages of genealogical research, Yakel [17] interviewed genealogists and observed their interactions during genealogy society meetings. Yakel found that genealogists' information seeking involves the gathering of factual information, which leads genealogists to connect and build a common identity with other genealogists and relatives. This process is open ended, as genealogists attempt to find more and more information about their ancestors. Yakel asserts that "family history should be viewed as an ongoing process of seeking meaning" and that genealogists' "ultimate need is not [to find] a fact or date, but to create a larger narrative, connect with others in the past and in the present, and to find coherence in one's own life" [17, Abstract]. The web, Yakel suggests, supports this process of connecting with others.

Yakel and Torres [18] expanded the findings from Yakel's 2004 study to describe group problem solving among genealogists at society meetings. At these meetings, experienced genealogists acted as an expert panel by fielding questions from the audience and collaboratively answering those questions. However, Yakel and Torres noted that members of the audience also offered advice and information to help the inquirer and the panel. Yakel and Torres also found that much of the advice and information offered by the panel and interviewees was instructional in nature. Yakel and Torres report that:

"The more experienced genealogists helped others by teaching them to interpret records, instructing them in search strategies and processes, and passing on other forms of both explicit and tacit knowledge such as the value of different types of records" [18, p. 98]

In a similar vein, Fulton [7] found that genealogists prize informal information sharing to advance their research interests. Social norms of genealogists include expectations of reciprocal information sharing and building relationships based on information expertise. Through her interviews with genealogists, Fulton identified a group of genealogists whom she called "super sharers"; these individuals spent considerable time helping other genealogists by providing instruction and advice to less experienced genealogists. Fulton also found that the web provided a means for genealogists to connect with family and fellow genealogists that went beyond the genealogists' immediate circles. She concluded that the web provided an important venue for information seeking and sharing among genealogists.

2.2 Online Question and Answer (Q&A) Communities

Much of the existing research on Q&A websites and user behavior seeks to characterize questions and answers to aid in the development of automatic methods of information retrieval and identification of good answers [e.g. 2, 3, 9, 10, 11, 21]. For example, Harper, et al. [9] developed a typology that divides questions into several types, including advice, approval/disapproval, factual, identification, prescriptive, and quality of an idea or concept. The authors applied this typology to three popular Q&A websites and found that factual (31%) and identification (28%) were the most common type of questions posted, while quality (7%) and disapproval (5%) were the least common. This dominant focus on automation, Shachaf and Rosenbaum [14] assert, has meant that the social interactions between users of Q&A forums have been largely neglected as an area of research. Rosenbaum and Shachaf further suggest that such studies "have not yet unpacked the black box of the

processes that characterize Q&A sites," and have not examined these sites "in the context of larger societal or even industry trends" [14, Introduction]. To these authors, Q&A websites have led to a new stage of question and answering that moves away from dyadic question negotiation to a collaborative approach to answering that is supported by an interactive computer technology.

What we do know about social interactions on Q&A websites is largely based on studies of large, multi-purpose Q&A websites, such as Yahoo! Answers and Answerbag. Drawing from these studies, several broad generalizations can be made about user behavior on multi-purpose Q&A websites, which may or may not be relevant to behavior on more specialized Q&A websites, such as Ancestry.com's message boards. In the Q&A environment, users take on different roles such as askers and answerers. Studies of large multipurpose Q&A websites, as well as a study of a programmer Q&A website (the Java Forum), have found that users who ask questions and users who answer questions form two distinct groups [15, 16, 20]. On these Q&A websites, there is consistently a group of answerers who respond to a disproportionately large share of the questions [15, 20, 1].

Gazen [8] takes this research a step further by categorizing answerers into two types, specialists or synthesists. Specialists are those who claim expertise in a given topic and answer questions without referencing other sources; synthesists are those who include one or more references to external sources in their answers. Analyzing user ratings of answers, Gazen found that synthesists' answers tended to be rated more highly by users than those answers provided by specialists. Nam, Ackerman and Adamic [12] found that higher levels of participation in Q&A forums correlates with better performance or higher quality answers.

Other researchers have explored the motivations for participation in Q&A communities. Yu, Jiang, and Chan [19] found reciprocity, reputation, enjoyment of helping others, self-protection, learning, moral obligation, and the advancement of the virtual community to be motivators. Constant, Sproull, and Kiesler [5] found altruism to be a strong motivator for answering questions, while strong social ties were not a strong motivator. Butler, Sproull, Kiesler, and Kraut [4] found that users participated to gather difficult to obtain information and to increase their visibility in social relationships.

3. METHODS

Existing literature on the information behaviors of genealogists provides an inchoate picture of how the web is shaping interactions among genealogists. Furthermore, there is a dearth of research that focuses on the social interactions among participants on Q&A websites dedicated to specific communities of practice. There is much left to learn about how media helps shape social interactions on the web. This study looks at both the social and technical dimensions of interactions on a website for a specific community of practice: genealogists. Questions that guided the research are:

- 1) What kinds of social interactions emerge among askers and answerers on a message board for genealogists?
- 2) In what ways do the mediated environment and existing practices of genealogists influence those social interactions?
- 3) Are Q&A forums serving as a means to educate users about the practices of a specific community of practice, in this case, genealogists?

To answer these questions, we examined interactions on Ancestry.com’s United States General Message Board. We extracted and coded text from all messages that appeared on this board during the period January 1, 2010 to December 31, 2010, totaling 1,086 posts. Institutional Review Board (IRB) approval was received for the use of these publically available data.

At the time of the study, the general message board had the largest number of posts on Ancestry.com, making it the most active forum on the site. Ancestry.com is one of the most popular genealogical websites in the United States, offering fee-based access to large databases containing public records data, such as census data and military service information. Other Ancestry.com databases, such as the family tree database, contain user-generated data. Annual fees for access to these databases are not trivial, ranging from 155 to 300 US dollars. Ancestry.com also offers ways for genealogists to connect to other genealogists, including its message boards. Unlike most of Ancestry.com’s website, the message boards are available free of charge to the general public. Anyone can use the message boards for genealogical research and the message boards currently contain over 17,000 million posts. Ancestry.com provides some oversight of its message boards by publishing message board use and etiquette guidelines on its website and providing mechanisms for users to report abuses of the boards. However, Ancestry.com states that it will not monitor the message boards for improper uses. As a result, the message boards appear to be self-defined and self-sustained by its users. Users of the studied Ancestry.com message board are employing the message board as Q&A forum. This also appears to be true for the other message boards on Ancestry.com.

Existing typologies taken from previous Q&A studies were used to develop an initial list of coding categories. The content of an initial set of two hundred posts was analyzed using the initial list of coding categories. Code refinement was carried out iteratively, with additional codes being identified and codes that did not adequately represent posts being replaced during the coding process. Coding was done at the post level; each post was given one or more codes based on its primary content. For example, an answerer post may be categorized as both factual and instruct if it includes both factual data and instructions on how to locate that data. We categorized all posts in each message thread, rather than only categorizing posts that appear to be “answers” or “questions.” Doing so allowed for the examination of all interactions between the posters, not only those interactions that have been deemed useful for automated retrieval (i.e. those that can be clearly defined as questions and answers). Message threads were also analyzed in light of what is known about the genealogical community of practice, the technological context, and the emergence of family history data sources on the web. A summary of the final coding categories is shown in Table 1. For a complete description of the specific categories that emerged from this analysis see Appendix A.

Table 1. Summary of categories

Askers’ Posts	Answerer’s Posts
Factual Question	Factual Answer
Instruct Question	Instruct Answer
Source Question	Source Answer
Request for Family Contact	Family Connection Answer
Unclear Question	Probes
Elaboration	Opinion
Gratitude	Encouragement

Factual answers were also examined to determine, if possible, the source of the data. In addition to coding individual posts, statistics were generated on the number and types of posts each asker and answerer generated. Other frequency data, such as the number of answers posted in response to a question was collected.

4. FINDINGS

4.1 Askers’ Posts

Between January 2010 and December 2010, 490 messages were posted on the general message board by individuals classified as “askers.” Of those posts, 191 were questions. Thirty percent of the asker posts were requests for factual genealogical data about specific families. Another 3% of the asker posts were questions about how to do genealogical research (instruct), 1.7% were requests for unknown living relatives to contact the asker, 0.9% were source questions (e.g. what website has marriage records?), and 4.2% of the questions were unclear as to what type of content the asker was requesting (see Figure 1). Other asker posts included gratitude (thanking the answerers) (32.8%) and elaboration posts (27.4%). Elaboration posts were those in which additional family information was offered in response to requests by the answerers for additional detail to aid them in the search process. The initial requests for factual genealogical data about specific families taken together with the elaboration responses account for 57.4% of the total asker posts.

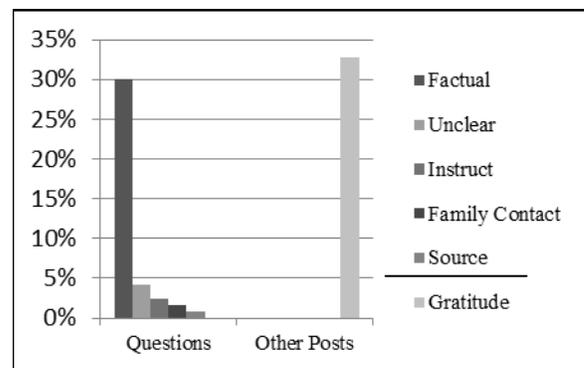


Figure 1. Percent of asker posts by content type

4.2 Answerers’ Posts

Between January 2010 and December 2010, 597 responses to questions were posted to the general message board by individuals classified as “answerers.” Factual answers were the most frequently provided type of content (58%), followed by requests for more detail from the asker (probes) (11.2%), instruct answers (11.2%), and source answers (10.7%). Answers that suggested a familial connection (4.4%), offered opinions (3.7%), or provided encouragement (.8%) were far less common (see Figure 2).

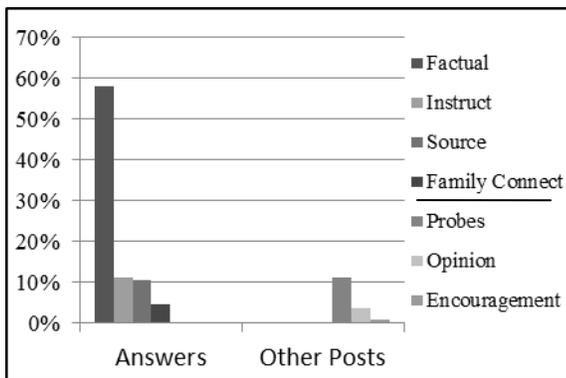


Figure 2. Percent of answerer posts by content type

Regarding the efficacy of the answerer probes, almost seventy percent (69.6%) of the answerer requests for more information (probes) received asker responses in which additional information was offered. In 4.9% of the factual answers, answerers copied and pasted data from Ancestry.com’s databases into their answers (the source of the data provided was indicated in the posts). In another 45.2% of the factual answers, the type of data given combined with the way the data was formatted suggests that these answers were also drawn from Ancestry.com databases. An additional 32.4% of the factual answers appeared to be copied and pasted from various other websites, as indicated by the answerer (e.g. “found this on the Internet”). In total, the researchers concluded that 82.5% of the factual answers contained data that were copied and pasted from various websites including Ancestry.com.

Of the 191 questions posted to the general message board, 13.6% received no responses. The questions that did not receive responses did not appear to differ significantly from those that had received responses, making it unclear why certain questions were answered and others were not. Of the 165 questions that received responses, 93.3% received one or more answer(s) that matched the question type (e.g., a factual question received a factual answer). On average, the response rate of answers to questions was 3.12 responses per query.

4.3 The Askers and the Answerers

One hundred and seventy-seven askers posted questions to the general message board. Of the askers, 98.2% posted one to two questions to the board, but a small number of askers (1.8%) asked three or more questions. Some askers posted more than one type of question (e.g. one factual and one reference question), so the total percentage shown for askers by question type exceeds 100% in Figure 3. The majority of the askers (77.2%) were looking for factual information about families, followed by 8.8% who posted unclear questions, 6.8% who posted instruct questions, 4.8% who posted source questions, and 2.4% who made family contact requests (see Figure 3).

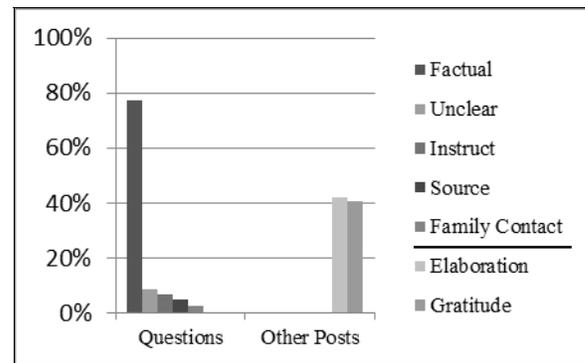


Figure 3. Percent of askers by content type

Of the 177 askers, 41.8% posted elaboration responses when answerers asked for more detailed information. Over forty percent (40.6%) of the askers thanked their answerers.

One hundred and twenty-four individuals provided answers. Many answerers posted more than one type of answer, so the total percentage shown for answerers by answer type exceeds 100% in Figure 4. The majority (53.6%) of answerers provided factual answers, followed by 24.8% who provided instruct answers, 16% who offered family connection posts, 15.2% who provided source answers, 8.8% who offered opinions, and 2.4% who provided encouragement.

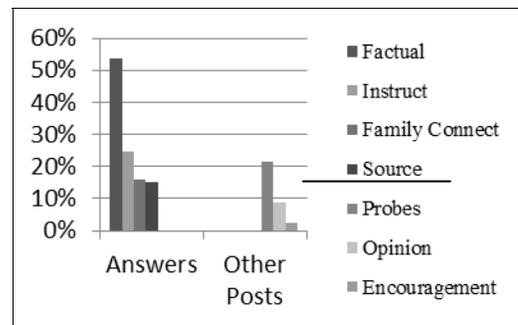


Figure 4. Percent of answerers by content type

Of the 165 questions that received responses, on average 2.1 answerers responded to each question. One hundred and thirty eight of the 165 questions that received responses had two or more answerers attending to each question. The majority (68.4%) of the answerers responded with at least one answer that fit the type of question asked. For example, answerers offered factual answers when askers asked for facts about their ancestors. However, 23.6% of the answerers preferred to give source or instruct answers to requests for factual data, rather than searching for the data themselves; in these cases, the answers provided links to specific websites that may have information pertaining to the posted question or gave instructions on how to find certain types of data. In cases where the asker did not provide enough information in the initial request, 21.6% of the answerers used probes to gather more data from the askers to aid in the search process.

4.4 Super Sharers

Over sixty-four percent (64.35%) of the answerers’ posts were provided by the top 10% of answerers. This finding suggests the presence of a small number of super sharers, who disproportionately provide answers to questions posted by other

genealogists. Of the super sharers, 66.6% used probes, while only 18.3% of the other 112 answerers used probes to gather more data from the askers.

5. DISCUSSION

5.1 The Nature of Genealogists' Interactions on Ancestry.com

We began with the questions “What kinds of social interactions emerge among askers and answerers on a message board for genealogists?” and “In what ways do the mediated environment and existing practices of genealogists influence those social interactions?”

Our findings suggest that the Ancestry.com message board is being used primarily by askers to locate factual data about specific families. Answerers, in turn, appear to be locating the requested factual data and copying and pasting that data into their answers. Although Yakel and Torres [18] found in face-to-face settings that genealogists provided assistance to other genealogists that was instructional in nature, most answerers in our study (53.6%) provided factual family history data, not instructional information (24.8%).

The high number of social interactions on the message board that involved asking for or providing family data may be the result of the accessibility of family history data on the web. The availability of factual data about families in a digital format on the web makes the provision of such data easier than if the data were only in a physical format. In turn, this accessibility may be shaping the nature of the online exchanges among genealogists on the studied message board. That is, given the ease with which genealogists can locate information and copy and paste that information in response to questions may decrease the likelihood that answerers will use instruct answers. Explaining how to accomplish something may be more difficult than simply copying and pasting the data into a message. Furthermore, askers appear to want factual data, not instruction, which may be a function of the perceived ease of access to the data.

Access to family data online and the ease to which this data can be copied and pasted into a message may facilitate the provision of factual answers, and may also be shaping other elements of the discourse between askers and answerers. The answerers' approach to “answering” questions by locating factual information about specific families for the askers may have encouraged the use of probes. Some answerers (21.6%) and many “super sharers” (66.6%) used probes to cull contextual data from the askers to help them locate the right data in the sea of genealogical information on the web. Genealogists, particularly seasoned ones, are aware of the difficulty of finding the right data. Many families share the same surnames, so finding the right data for the right family is difficult. Contextual information about the family becomes paramount when attempting to locate correct family data, particularly when searching large online genealogical databases in which hundreds, even thousands, of records exist for the same surnames. With such obstacles in place, many answerers wisely used probes to gain more knowledge of the family in question to improve the likelihood that the data they collected was correct. When answerers provide source answers or instruct answers, knowing specifics about families may not be as important in developing an answer. However, if answerers are attempting to locate highly specific data for the

askers, such knowledge is critical, and the use of probes becomes important in increasing the accuracy of their searches.

Answerer probes would be useless if the askers did not respond with further data. Almost seventy percent (69.6%) of the answerer requests for more information (probes) received asker responses (elaboration) in which additional information was offered. This indicates that askers do not always know what information to include in their questions and that answerers may benefit from using probes to extract information that was left unstated.

Yakel and Torres [18] found that access to records, specifically the lack of access, is the pivotal issue for genealogists. While increasing numbers of genealogical data are available online, many are only available for a fee. This includes a large portion of data that is contained within Ancestry.com's extensive databases. In 4.9% of the factual answers, answerers indicated that they copied and pasted data from Ancestry.com's databases into their answers. In another 45.2% of the factual answers, the type of data given combined with the way the data was formatted suggests that these answers were also drawn from Ancestry.com databases. The act of taking factual data about specific individuals directly from Ancestry.com and posting it on the message board is providing data to those who may not have a legitimate right to access that source, because they may have not paid for that access. However, Ancestry.com appears to turn a blind eye to this practice, possibly as a method of attracting new users or as a public relations tactic to retain current users, whom it suspects are sensitive about fee-based access to public records.

5.2 Online Exchanges and the Community of Practice

We hypothesized that the general message board might attract novices and may be serving as an entry point for newbies into the genealogical community of practice. This led to our third question: “Are Q&A forums serving as a means to educate users about the practices of a specific community of practice, in this case, genealogists?” We found that few askers stated they were beginners (4.5%). It also became clear that we could not assume that askers requesting basic family data were novices, as their requests may have grown from not having access to fee-based databases, rather than from not knowing how to find the data. Indeed, some askers specifically stated that they did not have access to Ancestry.com's databases and wanted someone “to do a look-up” for them. In the end, it was impossible to determine the experience level of the askers from the posts.

Answerers were likely uncertain about the experience level of the askers and whether askers needed instructional help. This uncertainty may have been one of the factors that influenced most answerers to provide factual data about families without including any “how to” information. Another more influential factor may have been that most askers requested information about particular family members. In turn, answerers may have felt obliged to offer such data, rather than instruct the askers on how to find that data. The ease to which family data can be copied and pasted from web-accessible resources may have further encouraged answerers to offer data rather than instruction.

Some answerers (24.8%), however, provided instructional answers. Many of these individuals provided instruct answers, even when the askers requested the provision of family data. This suggests that a small number of answerers were aware of their

role as teachers of genealogical research practices, not only their role as data providers. This group of answerers appeared to be knowledgeable genealogists, who were likely experienced practitioners. However, given the small number of instruct answers posted (11.2% of the answerers' posts), the message board may not be serving as an effective means of assimilating new genealogists into the community practice. Possibly, though, new ways of learning the genealogical craft are occurring on the message board. While not as directly pedagogical as instructional answers, answerer probes may serve as a means to augment the lack of instructional answers by educating askers about the kinds of information they need to know to be effective in locating the right family history data from web-accessible sources.

5.3 Cooperative Research, Not Collaborative Instruction

Yakel and Torres [18] found collaborative group problem solving behaviors in face-to-face meetings of genealogists. These group problem solving sessions allowed genealogists to share both implicit and explicit knowledge of the research process, and assimilate individuals into the community of practice. Group problem solving behaviors were also found on the studied message board, with over eighty percent (83.6%) of the message board threads revealing cooperative work among a group of answerers. One hundred and thirty eight of the 165 questions that received responses had two or more answerers attending to each question. Collaboration in which answerers jointly reached some sort of synthesis on an answer was rare on this message board. However, cooperation between answerers was common in that many answerers employed each other's answers to find additional data for the asker. The online environment and the availability of family data at one's finger tips may shape the types of cooperation that occurs on the message board, leading genealogists to engage in cooperative research, rather than collaborative instruction, as was found in face-to-face meetings of genealogists [18].

6. LIMITATIONS

This study explored interactions on only one of Ancestry.com's many publically available message boards, and, as a result, it may not represent the website's message boards as a whole. Additionally, it is impossible to know the actual intent of askers and answerers by examining their posts. Future research in this area is planned to include interviews with message board participants.

7. CONCLUSION

Overall, our findings suggest that the presence of online genealogical data and the affordances of interactive computer technologies appear to be pushing many answerers away from providing instruction on how and where to find family history data, which is common in face-to-face interaction, and pushing them toward providing those data outright. This in turn impacts cooperative work among answerers on the studied Q&A message board, leading answerers to engage in cooperative research, rather than collaborative instruction. However, new ways of providing "know-how" such as the use of answerer probes may be educating genealogists on the practice of genealogical research in the web context.

Our research also begins to address the need for exploring Q&A websites as communities with emergent practices by examining the social interactions between askers and answers on a message board that supports a well-defined group of practitioners. While some of our findings, such as the presence of super sharers, echo

the results of other Q&A studies, our findings extend this literature by introducing the concepts of answerer probes and asker elaboration.

8. REFERENCES

- [1] Adamic, L. A., Zhang, J., Bakshy, E., and Ackerman, M. S. 2008. Knowledge sharing and yahoo answers: Everyone knows something. In *Proceedings of the 17th International Conference on WWW* (Beijing, China, April 21 – 25, 2008). ACM, New York, NY, 665-674.
- [2] Bian, J., Liu, Y., Agichtein, E., and Zha, H. 2008. Finding the right facts in the crowd: Factoid question answering over social media. In *Proceedings of the International World Wide Web Conference* (Beijing, China, April 21 – 25, 2008). ACM, New York, NY, 467–476
- [3] Burger, J., Cardie, C., Chaudhri, V., Gaizauskas, R., Harabagiu, S., Israel, ... Strzalkowski, T., Voorhees, E., and Weischedel, R. 2001. Issues, tasks and program structures to roadmap research in question & answering (Q&A). DOI: <http://www-nlpir.nist.gov/projects/duc/roadmapping.html>.
- [4] Butler, B., Sproull, L., Kiesler, S., and Kraut, R. 2007. Community effort in online groups: Who does the work and why? In *Leadership at a distance*, S. Weisband, Ed. Taylor and Francis Group, New York, NY, 171–194.
- [5] Constant, D., Sproull, L., and Kiesler, S. 1996. The kindness of strangers: The usefulness of electronic weak ties for technical advice. *Organ Sci*, 7, 2, 119-135.
- [6] Duff, W. M. and Johnson, C. A. 2003. Where is the list with all the names? Information-seeking behavior of genealogists. *Am Archivist* 66, 79-95.
- [7] Fulton, C. 2009. Quid pro quo: Information sharing in leisure activities. *Libr Trends*, 57, 752-768
- [8] Gazan, R. 2006. Specialists and synthesists in a question answering community. In *Proceedings of the American Society for Information Science and Technology* (Austin, TX, November 03-08, 2006). 43, 1–10.
- [9] Harper, F.M., Weinberg, J., Logie, J., and Konstan, J.A. 2010. Question types in social Q&A sites. *First Monday*, 15, 7. DOI= <http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/2913/2571>
- [10] Hong, L. and Davison, B. D. 2009. A classification-based approach to question answering in discussion board. In *Proceedings of ACM SIGIR conference on Research and development in information retrieval* (Boston, MA, July 19-23, 2009). ACM, New York, NY, 171-178.
- [11] Jeon, J., Croft, W.B., and Lee, J. H. 2005. Finding similar questions in large question and answer archives. In *Proceedings of the 14th ACM International Conference on Information and Knowledge Management* (Bremen, Germany, October 31 - November 05, 2005). ACM, New York, NY, 84-90.
- [12] Nam, K.K., Ackerman, M.S., and Adamic, L. A. 2009. Questions in, knowledge in?: A study of Naver's question answering community. In *Proceedings of the 27th international conference on Human factors in computing systems CHI 09* (Boston, MA, April 4-9, 2009). ACM, New York, NY, 2357-2367.

- [13] Rainie, L., Lenhart, A., Fox, S., Spooner, T., and Horrigan, J. 2007. Tracking online life. Pew Internet and American Life Project. DOI=
http://www.pewinternet.org/Reports/2000/Tracking-Online-Life.aspx
- [14] Shachaf, P. and Rosenbaum, H. 2009. Online social reference: a research agenda through a STIN framework. Paper presented at the iConference, (Chapel Hill, NC, February 08-11, 2009). ACM, New York, NY. DOI=
http://ischools.org/images/iConferences/shachaf-rosenbaum_iconf091.pdf
- [15] Shah, C., Oh, J.S., and Oh, S. 2008. Exploring characteristics and effects of user participation in online social Q&A sites. *First Monday*, 13, 9. DOI=
http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2182/2028
- [16] Welsler, H.T., Gleave, E., Fisher, D., and Smith, M. 2007. Visualizing the signatures of social roles in online discussion groups. *Journal of Social Structure*, 8, 2, 1-32.
- [17] Yakel, E. 2004. Seeking information, seeking connections, seeking meaning: Genealogists and family historians. *Inform Res*, 10, paper 205.
- [18] Yakel, E. and Torres, D.A. 2007. Genealogists as a “community of records.” *Am Archivist* 70, 93-113. Welsler, H.T., Gleave, E., Fisher, D., and Smith, M. 2007. Visualizing the signatures of social roles in online discussion groups. *Journal of Social Structure*, 8, 2, 1-32.
- [19] Yu, J., Jiang, Z., and Chan, H. C. (2007). Knowledge contribution in problem solving virtual communities: the mediating role of individual motivations. In *Proceedings of the 2007 ACM SIGMIS CPR conference on Computer Personnel Research* (Montréal, Canada, December 09 - 12, 2007). ACM, New York, NY, 144–152.
- [20] Zhang, J., Ackerman, M. S., and Adamic, L. 2007. Expertise networks in online communities: structure and algorithms. In *Proceedings of the 16th international conference on World Wide Web* (Banff, Canada, May 08-12, 2007). ACM, New York, NY, 221-230.
- [21] Zhang, Y. and Fu, W.T. 2011. Designing consumer Health Information Systems: What Do User-Generated Questions Tell Us? In *Foundations of Augmented Cognition, Directing the Future of Adaptive Systems, Lecture Notes in Computer Science*, Volume 6780/2011, 536-545. Springer-Verlag, Berlin, Germany.

APPENDIX

A. MESSAGE CATEGORIES

Askers' Posts		
Code	Description	Example
Factual Question	Request for genealogical factual data from sources, such as census, birth, death, or marriage records.	Can anybody give me any information about the above who I believe may have moved to the United States...I don't have Ancestry Worldwide so can't search the records for the US.
Instruct Question	Asking for instruction on how to conduct some aspect of genealogical research, such as how to interpret historical records, how to conduct successful searches in online databases, and how to navigate governmental systems to obtain needed documents.	I think that I still do not know how to use ancestry.com search feature? Could help please? Many thanks for the link to the US National Archives. It is a maze of info, and I am not sure how to use? So how do I go about reconstructing his visit to the US? Is there any hope of doing so?
Source Question	Asking where specific information is located.	Where can I get naturalization records?
Request for Family Contact Question	Requesting living relatives to contact them.	If you are descendants of de van Brederode I want to ask you if you want to contact me?
Unclear Question	Purpose of the question is unclear. Often data given without clear request.	Hi, I am researching my great-grandmother Annie, who was born 1870 in Yorkshire, England.
Gratitude	Asker thanking answerers for their help and the information provided.	Thank you so much!!
Elaboration	Asker offering additional information (beyond what was offered in initial query) to aid the answerer in their search; often offered in response to answerer's request for additional information/data.	My Johnson came from Sweden and was born 1855.

Answerers' Posts		
Code	Description	Example
Factual Answer	Factual data drawn from existing sources such as census, military, or marriage records.	US/Canada border crossings Name: Harriet Chance Arrival Date: 29 Nov 1917 Age: 34 Birth Date: abt 1883
Instruct Answer	Instructions on how to conduct genealogical research, such as how to interpret historical records, how to conduct successful searches in online databases, and how to navigate governmental systems to obtain needed documents.	If she was adopted, it's likely she would have had some biological relationship with the family. I would examine all the baptism records of her children, and all her (supposed) siblings, noting the godparents. You might find the answer to your puzzle in the records of the extended family.
Source Answer	Referring the user to particular sources for genealogical data, such as specific websites.	check the Ellis Island site for immigration records - http://www.ellisland.org
Family Connection	Answerers suggesting that they are related to the askers.	Edmund was my step-grandfather...He died in 1962, when I was still in college.
Probes	Answerers requesting more information from the askers to help them conduct research for askers.	Do you have clues as to Ida or her family which might help in a search? Birth date? Date of emigration?
Encouragement	Offering encouragement to the asker as they go forward in their research.	Good luck in your search!
Opinion	Voicing subjective ideas.	I think that the Ellis Island website could be improved.