Reexamining Technological Support for Genealogy Research, Collaboration, and Education

FEI SHAN, Virginia Polytechnic Institute and State University, USA
KURT LUTHER, Virginia Polytechnic Institute and State University, USA

Genealogy, the study of family history and lineage, has seen tremendous growth over the past decade, fueled by technological advances such as home DNA testing and mass digitization of historical records. However, HCI research on genealogy practices is nascent, with the most recent major studies predating this transformation. In this paper, we present a qualitative study of the current state of technological support for genealogy research, collaboration, and education. Through semi-structured interviews with 20 genealogists with diverse expertise, we report on current practices, challenges, and success stories around how genealogists conduct research, collaborate, and learn skills. We contrast the experiences of amateurs and experts, describe the emerging importance of standardization and professionalization of the field, and stress the critical role of computer systems in genealogy education. We bridge studies of sensemaking and information literacy through this empirical study on genealogy research practices, and conclude by discussing how genealogy presents a unique perspective through which to study collective sensemaking and education in online communities.

CCS Concepts: • **Human-centered computing** → **Empirical studies in HCI**; **Empirical studies in collaborative and social computing**.

Additional Key Words and Phrases: genealogy, family history, community of practice, collaborative sensemaking, information literacy, andragogy, online education

ACM Reference Format:
Fei Shan and Kurt Luther. 2025. Reexamining Technological Support for Genealogy Research, Collaboration, and Education. *Proc. ACM Hum.-Comput. Interact.* 9, 2, Article CSCW155 (April 2025), 33 pages. https://doi.org/10.1145/3711053

1 Introduction

Genealogy, the activity of studying family history and lineage, has a long history itself. Genealogists search for information regarding ancestors among surviving documents, discover history of a family through traces of individual family members, and organize articles and pedigree diagrams indicating parental relations of the family. Emerging genealogy websites such as Ancestry.com and FamilySearch.org has transformed traditional genealogy research since the early 2000s. These sites granted users the capability to access millions of digitized historical documents through the Internet, which attracted an increasing number of users who are interested in exploring histories of their families. Featuring the idea of discovering common ancestors and making connections to living relatives around the world, genealogy websites also allowed users to build family trees and connect them to others' trees, presenting a unique form of massive-scale online collaboration [20].

These dramatic changes reverberated throughout the genealogy community, eliciting mixed reactions. In 1999, Elizabeth Shown Mills, president of the American Society of Genealogists and editor of *National Genealogical Society Quarterly*, raised concerns that "family history's progress

Authors' Contact Information: Fei Shan, fshan@vt.edu, Virginia Polytechnic Institute and State University, Virginia, USA; Kurt Luther, kluther@vt.edu, Virginia Polytechnic Institute and State University, Virginia, USA.

This work is licensed under a Creative Commons Attribution 4.0 International License.
© 2025 Copyright held by the owner/author(s).
ACM 2573-0142/2025/4-ARTCSCW155
https://doi.org/10.1145/3711053

as an intellectually valuable pursuit" could be undermined by "web-grown genealogists" who were "largely unschooled in research principles" but empowered by technologies to broadcast their genealogy work [73]. To ensure the quality of the genealogical research product, senior members of this community sought to formalize the desired research processes into guidelines and standards [27, 98]. However, these standards failed to gain traction outside of professional societies, and concerns of widespread genealogy research errors proved prescient.

Indeed, as the spread of misinformation became a global concern, the genealogist community has faced its own version of this problem. A 2014 study [116] suggested that there was a considerable number of genealogy hobbyists, referred by serious genealogists as "clickologists" and "half-researchers", who built and connected to family trees on genealogy websites without knowing the proper genealogy research practices. Such behavior brought disturbances to the genealogist community: People found their close relative became a member of many strangers' families, and "research hints," which are algorithmic recommendations of relevant information about an ancestor, would pick up inaccurate information from family trees and mislead more users. In recent years, an explosion of interest in genealogy has only intensified the problem. As of 2019, FamilySearch claimed to be the home of the world's largest family tree, built collectively by its users and containing more than 1.2 billion ancestors [76].

Surprisingly, there has been minimal follow-up research from an human-computer interaction (HCI) or CSCW perspective on this challenge in the past decade. Recognizing the extensive growth of genealogical business fueled by the advancement of new technologies such as customized search engine [48], genealogical data visualization and analysis [10, 14, 67, 108], DNA testing [15, 37, 78, 89, 100], and artificial intelligence (AI) based handwriting recognition [9, 52, 82], face recognition [74, 90] and family tree network exploration [69, 70], we believe there is an urgency to reexamine the current state of technology-supported genealogy research. In this paper, we attempt to fill the research gap by conducting a qualitative study of current genealogy practices. We extend prior work with new analysis comparing amateur and expert genealogists, investigating the impact of new technologies, and understanding how newcomers of this community learn to become competent researchers on the Internet. Specifically, our study is guided by the following research questions:

(1) How has genealogists' research evolved with recent technology advances?
(2) What are amateur genealogists' research practices compared with established standards?
(3) How do amateur genealogists learn online, and how can technology better support this process?

We conducted a semi-structured interview study with 20 genealogists, recruiting 10 experts with extensive professional research experience and 10 amateurs who are relatively new to this field. To the best of our knowledge, this study is the first that considers the role of standardization, professionalism, and credentials in the genealogist community, and to conduct a side-by-side comparison of research practices between genealogists with different levels of expertise. This work presents a more comprehensive description of genealogists' research processes and reveals several concerning phenomena. We also propose the design recommendations based on our findings, and discuss the potential of genealogy as a unique lens through which to study a range of important problems in the field of HCI.

2 Related Work

2.1 The Genealogy Research Process

Academic studies of genealogy research practices and community have been scattered across various disciplines, including information science [33, 71, 119], library and archival science [16, 23],

sociology [11, 61, 64], tourism [53, 88, 103], and public history [24, 39]. In 2003, Duff and Johnson [23] presented a pioneering user study focusing on genealogists' information seeking behavior with archival systems. Their work revealed the complexity of genealogical research, which is more than just "collecting names and building family trees." They found that genealogists need to refer to historical context and apply searching strategies in their research. As archival systems did not meet their needs, genealogists developed their own finding aids, consulted others who had knowledge about a period, geographic area, or community, and formed a strong network through "courses, genealogical society meetings, and newsletters" [23].

Duff and Johnson [23] also pointed out that the genealogical research is an iterative process; experienced genealogists have developed a set of strategies which they can apply each time they look for new information. Similarly, Francis [30] found a "circular pattern" in genealogy research. The information genealogists seek might not be available or easily found; therefore, genealogists often need to circle back to previous steps of their research and look into other research directions. Friday [32] summarized these observations and proposed a "genealogical fact" frame [31] — describing genealogical evidence with four elements: name, date, place, event — to establish an iterative genealogical search model. Darby and Clough [19], at a higher level, suggested that genealogists would start with ancestors who left abundant traces in easily accessible records, then push backward in the timeline to those who are more difficult to research.

Yakel [119] and Fulton [35] revealed aspects of genealogical research beyond information seeking, which also involves documenting, organizing, producing, and sharing with family members. Intrigued by the story of ancestors, many of Yakel [119]'s interviewees perceived themselves as narrators, archivists, and navigators of their family history. Veale [104], Fulton [33], Bishop [8], and Moore and Rosenthal [75] also elaborated the motivations of genealogists. These studies, through different methods and scopes, characterized genealogists as enthusiasts, explorers and investigators driven by an obligation to preserve family history and motivated by self-exploration through the stories of ancestors. They enjoyed the process of investigation and resolving genealogical challenges, and devoted a large portion of their free time to this hobby.

There are also genealogists who are deemed experts in this field, who have demonstrated their proficiency in genealogy research and provide professional services for individual clients and government agencies. The National Genealogical Society (NGS) in the U.S. recognizes two sources of genealogy credentials: the certification granted by the Board for Certification of Genealogists (BCG) in Washington, D.C., and the accreditation by the International Commission for the Accreditation of Professional Genealogists (ICAPGen) in Orem, Utah [97]. Genealogists may earn these credentials by submitting their assessment and documented research project as evidence of their knowledge and research competence [26, 28]. However, these formal indicators of expertise have not been taken into consideration in previous studies. Instead, study participants have been described as "professional" [23], "non-professional" [75], "amateur" [33–35, 75], "hobbyist" [35, 105], or by the number of years they have spent in genealogical research [45, 92, 119, 120]. As a result, we cannot compare genealogists' research practices except for occasional hints from authors. For instance, Duff and Johnson [23] noted that novices might have less knowledge of history background and experienced difficulty converting their need of information to queries of records through archival systems. As other examples, Darby and Clough [19] suggested the "easy tree-build" stage required "little expert knowledge", and Fulton [33]'s respondents would not mind hiring professional genealogists for efficiency, but preferred conducting research by themselves for enjoyment. These findings suggest that amateur genealogists might be less efficient in research, but how their research practices might differ from experts' is unknown, and the quality of participants' research work was not assessed.

From an HCI perspective, Willever-Farr and Forte [116] were the first to investigate the quality and collaboration of genealogy research on the Internet. They conducted a study of two user-contributed genealogy websites, Ancestry.com and FindaGrave.com (since acquired by Ancestry), and found that the erroneous user production was a major cause of misinformation and conflicts on these websites. Furthermore, participants of their study suggested that Ancestry's advertising strategy and "research hint" feature attracted users with little knowledge of the genealogical research process and encouraged them to produce pedigrees rapidly without concern for proper verification. The presence of careless contributors and inaccurate or even fabricated information undermined the overall reliability of user-generated content, introduced burdens to other genealogy researchers, and hindered collaborations on these platforms. However, we still do not have a clear view on how novice genealogists' mistakes are generated in their research, or how technological support may help or hinder the problem.

2.2 Sensemaking Challenges and Information Literacy

Genealogical research can be seen as a collaborative sensemaking process [56]. With a broader definition, sensemaking covers a wide range of human information behavior [118], spanning information foraging, synthesis, production, and more [62, 87, 117, 122]. As sensemaking pitfalls and challenges have not been explicitly studied in the context of genealogical research, we look for inspiration from other domains.

Intelligence analysis, investigation, and data analytics are common research fields of human sensemaking [46, 59, 85]. Heuer [46] summarized how cognitive biases and limitations of memory and perception may lead to analytic pitfalls in intelligence agencies. Studies on criminal investigation pointed out factors that may undermine judgement of investigators, such as social norms on efficiency in the work environment [3] and asymmetrical skepticism [72]. Meanwhile, studies of data science found that the lack of clear structure and contextual information prevent users' form understanding and reusing data effectively [25, 38, 59, 123]. These studies often focus on the work of trained professionals, whereas in the genealogist community, a considerable amount of online genealogical work is conducted by amateur hobbyists, who might face more challenges that would not be fully captured in these settings.

Studies of collective sensemaking take into account social context and interactions between individuals [21] and surface more complexities. For instance, Kang et al. [50] proposed the concept of "teammate inaccuracy blindness" to describe one's unawareness of information quality provided by other team members. Koesten et al. [60] suggested that when people indeed need to assess data, effective communication with the data producer and other users is beneficial but not well supported. Balakrishnan et al. [5] argued that sharing all information between team members at once may enhance initial confirmation biases. Finally, Li et al. [66] analyzed how people made errors depending on the given context at each stage, which then propagated across the sensemaking pipeline. Another paradigm of crowdsourced sensemaking research focuses on a bottom-up, novice-led environment [6], in which the quality of collective sensemaking depends more on individual capabilities, and pitfalls could lead to negative consequences such as misinformation and digilantism [22, 47, 81, 94, 101]. We have seen deployment and evaluation of crowdsourcing approaches for related tasks such as transcribing and analyzing records [41, 52, 107] and identifying historical photos [74], but not one directly applied to genealogy research.

As genealogy research comprises extensive use of archives and libraries, another research field closely related is information literacy, a theoretical framework describing an individual's capability to "seek, evaluate, use and create information" [102]. This concept was originated from the library and information science discipline. Though traditionally focused on print culture, it evolved in past years — associated with the concept of digital and media literacy, which emphasize

aspects working with digital information and computer — to address information challenges today [7, 68, 83]. Information literacy skills were recognized as a critical component in combating online misinformation [49, 54, 95]. In the genealogy domain, librarians and archivists focused on educating genealogy beginners frequently encounter audiences with limited information and digital literacy skills [12, 114]. They reflected on amateurs' insufficiency with computer systems and accessing online information, underlining the importance of investigating how amateurs' skills could be improved in a technology-intensive genealogy research context.

2.3 Learning in Online Communities

Genealogy was depicted as life-long learning endeavor through which practitioners learn about family history, themselves, and proper genealogy research methods [45]. In this process, amateur genealogists might skip initial learning on how to conduct genealogy research, diving in with a trial-and-error approach [19]. Genealogical education resources are presented in various forms. Veale [105] surveyed online educational content for genealogists and categorized it in three groups — *scholarly*, *topical*, and *ad-hoc* — arguing that the *scholarly* education was more formal, theoretical, and authoritative, while *ad-hoc* resources had its strengths in personalization, problem-based learning, scale, and immediate availability. Veale [105] further concluded that genealogists' learning experiences had yet to fully take advantage of a digital, online environment. Participants in Willever-Farr and Forte [116]'s study echoed this observation, mentioning that they were willing to help inexperienced genealogists, but the online environment did not provide sufficient support for their educational activities.

Learning is not only the acquisition of abstract knowledge, but also an cultural activity comprising rich social interactions [79]. Legitimate peripheral participation (LPP), as the core concept of situated learning, describes the process of newcomers to a community starting with participating in peripheral work and gaining mastery along the way [65]. An example that has been extensively studied by HCI researchers is the Wikipedia community [13, 106]. However, Wikipedia enforces a "no original research" policy, and edits on the site needs to be reviewed and approved by senior members of the community, which essentially constrains community members' use of information. In contrast, in the genealogist community, it appears that anyone has full permission to conduct original research on their family and share research products freely. It complicates the community practice and demands more careful learning experience design for the genealogist community.

Newcomers are also expected to develop their identity and acceptance of the values and norms of a community of practice (CoP) through participation [17, 40, 109]. Whereas various studies suggest that conducting genealogy research enhances people's self-identification [43, 75, 96], Handley et al. [40] raised concerns that the intention of preserving "sense of self" may undermine effectiveness of learning. Furthermore, Roued et al. [92] found that genealogists often move among multiple genealogical and general-purpose platforms, which may pose more challenges in conveying community norms. Overall, it is unclear from online learning and HCI scholarship how (or whether) amateurs would adopt established practices as members of the genealogist community as they pursue their genealogy research.

3 Method

3.1 Recruitment

To understand the perspectives of genealogists with diverse levels of competence, we recruited participants through two channels: (1) directly inviting professional genealogists via email, and (2) posting recruiting ads on an open online genealogy forum.

For the first channel, we leveraged the member directories of the two major credentialing organizations for genealogical research in the U.S.: BCG and ICAPGen. As of August 2023, the two organizations have 257 and 217 certified/accredited members, respectively, with a few of them obtaining both credentials. We sent invitations with an information sheet of our study to candidates' email addresses obtained from the organization directories. Respondents who expressed their interest in participating were then asked to fill out a survey questionnaire and confirm their intent of participating before we scheduled the interview.

We also recruited from the public mainly for reaching out to genealogists who are relatively new to this field. We posted recruiting information on the r/Genealogy subreddit, the largest genealogy-purposed public online forum we identified with more than 100,000 members. The post contains a brief summary of our study with links to the information sheet and survey questionnaire. We then contacted survey respondents who expressed interest in participating through the email address they provided to schedule the interview.

In the recruitment survey, genealogists first provided their basic demographic information like gender, age range, and ethnicity. We then asked them about their genealogy research experience, including how long they have been conducting genealogy research, whether they research on their own family, others' family, or for clients as a paid professional. We also asked if they have received formal education/training in genealogical research, if they joined any genealogy related groups, organizations, or online communities, and if they hold any genealogy related credentials. Furthermore, to obtain a preliminary understanding of the research tools and technologies our participants would utilize, we let respondents provide a brief summary of the tools, software and/or websites they would typically use when they were conducting genealogy research. Lastly, we asked respondents to review the attached information sheet and confirm their interest in participating in the interview. The information sheet stated the purpose and procedure of the interview study, participants' compensation for completing the interview ($20), and that the study is approved by the Institutional Review Board (IRB). A couple of our interviewees generously declined the offered compensation.

3.2 Interview Design

We conducted in-depth, semi-structured interviews with participants. As technology-supported genealogy research practices can be multifaceted, our interview questions were organized into several sections. The introductory questions covered the background research experience of our participants, extending some of the questions in the survey. The rest were designed based on the high-level research questions described in Section 1. Inspired by literature from this community and academia, we looked into several aspects of genealogy research: (1) How genealogists handle research that lasts a long time without reaching a definitive conclusion; (2) How they overcome research barriers, commonly known as "brick walls", (3) How they perceive genealogical information and resolve unreliable information, and (4) How they share research findings to others. In each of these aspects, we asked participants how technology has helped or hindered their research. We then asked about their learning experiences. Finally, participants were invited to bring up important aspects about their research that we did not cover. A detailed interview guide is provided in Appendix A.

3.3 Data Collection and Analysis

The recruiting survey responses were collected with Google Forms. Twenty out of 56 respondents participated in our interview study. Half of the interview participants were experienced genealogists who conduct professional research, nine of whom hold genealogy credentials and are recruited from direct invitation, and the other came from public recruiting. This group of participants shared

ID	Gender	Age	Years of Experience	Genealogy Education
P1E	F	41–50	36–40	Self-taught
P2E	M	71–80	21–25	Self-taught
P3E	F	41–50	41–45	College degree related to genealogy
P4E	F	18–30	6–10	College degree related to genealogy
P5E	F	61–70	11–15	Week-long lectures, workshops, genealogical society courses
P6E	M	61–70	46–50	Study group, conferences
P7E	M	61–70	41–45	Certificate program, genealogical institute programs, study group
P8E	M	31–40	6–10	College courses, genealogical institute program
P9E	M	31–40	26–30	College courses, self-taught
P10E	N	71–80	46–50	Self-taught
P11A	F	41–50	1	Self-taught
P12A	F	18–30	2	Self-taught
P13A	F	18–30	5	Self-taught
P14A	F	31–40	1	Self-taught
P15A	F	31–40	1	Self-taught
P16A	M	31–40	3	Self-taught
P17A	M	61–70	4	Self-taught
P18A	M	18–30	2	Self-taught
P19A	M	41–50	2	Self-taught
P20A	F	51–60	2.5	Self-taught

Table 1. Demographic information about our interview participants. The ending letter of each alias indicates the experience level of the participant: A — amateur, E — expert. Gender: F — female, M — male, N — not disclosed.

with us their rich experience researching for individual clients, teaching lectures, publishing and editing family history books and journals, working for major genealogy companies, organizations, libraries and government's public repatriation programs, and taking leading positions in local and nationwide genealogical societies. In our later analysis, we refer to them as *experts*. The other half of participants were considerably newer to this field. They had no more than 5 years of experience, did not receive formal education or training, nor did they have professional genealogy research experience. They are referred to as *amateurs* in our following discussion.

Overall, our participants are adults between 18 and 80 years old, with a median age range of 41–50. Participants self-reported as White, Hispanic or Latino, Black or African American, and American Indian or Alaska Native ethnicity, with the majority being White. Participants' gender distribution is balanced across the expert and amateur groups. Table 1 provides a general description of our participants' research and educational experience. For ease of reading, a letter is appended to each participant's pseudonym as an indication of the his/her expertise, where *E* stands for *expert* and *A* stands for *amateur*.

All our interviews were conducted remotely and recorded with Zoom software. Most of the interviews took about one hour, with a few exceptions extending to 90 minutes. Transcripts of the interviews were auto-generated by Zoom and verified and analyzed manually. The transcripts were anonymized and coded using the Google Workspace package. Specifically, we adopted thematic analysis [18] to iteratively extract themes from the interview transcripts. These themes were then divided into three high-level topics, associated with either genealogy research practices

(Section 4.1 to Section 4.5), interactions within the community (Section 4.6), or learning and education in genealogy research (Section 4.7). Our coding themes are presented as the headings of sub-subsections in Section 4 and summarized in Table 2. For each theme, we discuss how amateurs' approach may differ from that of experts and the challenges genealogists are facing during their research, which are emphasized in **bold** text.

4 Findings

Throughout our study, the Genealogical Proof Standard (GPS) [91] was frequently referred to by experts to describe their genealogy research process. The Board for Certification of Genealogists (BCG) summarizes GPS with the following five components [27] :

(1) Reasonably exhaustive research;
(2) Complete and accurate source citations;
(3) Thorough analysis and correlation;
(4) Resolution of conflicting evidence;
(5) Soundly written conclusion based on the strongest evidence.

We also found that amateurs occasionally follow similar research practices with experts, applying some aspects of the GPS implicitly without realizing it. Therefore, we adopted GPS as a high-level framework in our later iterations of thematic analysis. We also leverage this framework when discussing genealogists' research process, categorizing our coding themes based on how they contribute to each component of GPS from Section 4.1 to Section 4.5.

It is important to note that the way we present our findings as experts' versus amateurs' approach does not imply that experts' research is strictly superior. Our participants emphasized that there are hobbyists who do not have credentials but are skilled at genealogy research, and even professional genealogists could make mistakes if they are careless or researching an area that is unfamiliar to them. That being said, we believe it is still important to identify established genealogy research practices realized by expert genealogists, and undesirable ones that are commonly presented in amateurs' research.

Following the findings of genealogists' research practice from GPS, we investigate social context of the community and current learning and educational approaches and constraints, aiming to build a comprehensive view on how amateurs' research and learning can be supported in this community.

4.1 GPS Point 1: Reasonably Exhaustive Research

Collecting information is the foundation of genealogy research. First and foremost, expert genealogists expect an *reasonably exhaustive* search of information before reaching a conclusion. This requirement has multiple implications for genealogists' research practices.

4.1.1 Knowledge of records: what's out there, what information is contained, how to access them. An exhaustive search of information requires genealogists to obtain solid knowledge about the availability of historical documents, including what type of records still exist, what information is contained in each type of records in a certain era, and how to access these records. Expert genealogists demonstrated such knowledge throughout our interviews. As one expert emphasized, "You have to know where the records are. So that's really the first step" (P5E).

Such knowledge appears not only helps experts search relevant information exhaustively, but also allows them to obtain an overview of the research scope beforehand and plan their research accordingly:

> We would generally have an idea of where we want to go and how we wanted to spend the time before we even spent it. We weren't just going around. We knew from the

Category	Coding Themes
GPS Point 1: Reasonably Exhaustive Research	• Knowledge of records: what's out there, what information is contained, how to access them • Traversing online and local records • Focusing on a concise question • Strategies for overcoming brick walls • DNA as new evidence
GPS Point 2: Complete and Accurate Source Citations	• Organizing with file systems and family trees • Keeping track with research logs
GPS Point 3: Thorough Analysis and Correlation	• Knowledge of information production • Verifying clues by oneself
GPS Point 4: Resolution of Conflicting Evidence	• Keeping an open mind dealing with uncertainties and new information • Resolving conflicts with genealogy timelines • Testing hypotheses with family trees
GPS Point 5: Soundly Written Conclusion	• Producing research reports and family tree data
Community Interactions	• Privacy concerns in online collaboration • Peer review and conflicts • Perceptions of expertise
Genealogy Learning and Education	• Learning by doing • Seeking the right learning resources • Knowing without implementing • Learning with genealogy software systems

Table 2. Summary of findings.

> get-go, based on the problem, what record sets we would need to check and how long approximately it would take us to go through a certain amount of records. — P8E

In comparison, amateurs can be **less knowledgeable about the records**, according to experts, which risks their information search being incomplete or inefficient. Indeed, our amateur participants mentioned their unsuccessful search when they did not know records well enough:

> There was a fire, and a ton of records were destroyed. Well, I spent a long time trying to find this information, and it literally just didn't exist anymore. And I don't think I realized that until I had already spent a lot of time looking for it. — P15A

4.1.2 Traversing online and local records. The emergence of genealogy websites and digitized records in recent decades grants a new level of convenience to genealogists and greatly accelerates their research. One expert described the transformation of this field as a "massive digital renaissance in genealogy" (P4E). Expert and amateur genealogists expressed great appreciation for having access to digitized historical documents on the Internet:

> The computer has made my work so much easier. I can do in an hour what it would have taken me a week to do in the old days. Because I can sit here online even with my crummy Internet, and I can get sources that I would have had to travel hundreds of miles, spent days in a courthouse or in some archives with a pencil and paper doing which I can now do almost instantaneously. — P10E

Along with the record digitization, documents will be transcribed and indexed to support refined searches of data entries with keywords and other metadata. Algorithmic research hints will suggest records and user-built family trees to the users on genealogy websites. Notably, some expert genealogists are skeptical to the inaccurate information that research hints and family trees contain, but they do not object to leveraging these features in their research. As P5A argued, "There might be a kernel of truth in there." Another expert genealogist felt it is necessary to get a sense of the work that has been done by other researchers presented in online family trees:

> If you haven't done anything and you just want to see what's out there, it makes sense that you would look at the hints that you would look at online trees. You want to see what other people have done, so that you're not recreating everything. — P3E

Genealogists are likely to visit different genealogy websites to access unique record collections and maximize their search coverage. They mentioned even searching through the same record collection on multiple websites to leverage the variation of search algorithms. Still, online public documents is far from the entirety of genealogists' information source. Despite collecting online records, many genealogists suggested that there are reams of local documents not digitized or easily accessible to them, and one has to reach out to many kinds of record keeping facilities like government offices, archives, churches, history and genealogy societies, etc., to acquire these records. Furthermore, expert genealogists suggested that living family members who possess unique private information, such as old family photos, diaries, or even stories, is another viable source of information that genealogists should consult at the beginning of their research:

> I want to get a very clear assessment [on] what they know about that ancestor, what they already have gathered about them as far as historical documentation, or what some of the family stories, perhaps, are. And then I like to make sure that there isn't another close family member who may know more than them that we need to potentially interview or connect with. — P9E

Participants complained that there are amateurs who **stick to a few sources** on genealogy websites, who "sit at Ancestry," "think all the records are online," and have "limited exposure" to obscure, local records. Compared with experts, they are likely not be able to satisfy the reasonably exhaustiveness requirement of GPS. Genealogy websites might have a negative impact on this matter, as a few participants (mainly amateurs) suggested that some websites are "consumer products" that have financial incentives to monopolize the record collections and keep users within their platform.

4.1.3 Focusing on a concise question. Exhaustive search implies the possibility of collecting massive amounts of information for genealogists to process. An ancestor may leave their life traces in all types of documents, and each document contains various information pieces which could lead to new search directions. Therefore, expert genealogists felt that conducting reasonably exhaustive research without a concise research question is impractical. They proposed that genealogists should form concise research questions and try to focus on one question at a time. Furthermore, one participant elaborated that concise questions help genealogists to break down their research into manageable sections:

> One of the things that that many people don't understand is that, especially when you're doing this as a professional, you really want a very concise question . . . if you don't have a way to say "I'm done," research goes on forever. . . . You really want a tight concise question, and it's better if it's a yes or no. — P6E

On the other hand, multiple amateurs mentioned that they often **lose focus** in the ocean of documents without a concise research question, and soon be overwhelmed by the information

presented to them. Additionally, it appeared that automated research hints might worsen this distraction problem for amateurs:

> There's something called [a] rabbit hole, where you find a little hint and then all of a sudden your main goal for that day is pushed off to the side and you're chasing a rabbit down this other hole over here to see if it comes up with anything for you. And I find that to be quite the case . . . A lot of documents are getting online and indexed and so sometimes your initial goal for the day is just totally blown away. — P20A

4.1.4 Strategies for overcoming brick walls. One important aspect of genealogy research is that genealogists may not always find the direct answer to their question immediately. When hitting a so-called "brick wall," expert genealogists proposed several approaches to overcome it. First, they suggested that re-researching from a different perspective of the subject or looking for leads from relatives around them may help genealogists to find new research leads:

> Couple of times I've hit the wall . . . and I just put it away for a few days and then start from a different place to try to find it. You know, start from a different family member or a cousin or somewhere else, just completely different than where I was, to see if I can have a breakthrough. — P2E

Second, expert genealogists may advance their approach when there is a big chunk of missing information. Instead of focusing on a particular ancestor, they would research a group of people who share common life experience, such as being members of a family or ethnic community, which suggests shared living patterns among them. Expert genealogists call this approach *cluster research* and suggested it was a more advanced research method not well known to amateurs.

> There was a gap there of almost a whole generation. You had to sort of use extraneous clues and look for groups of people together. . . . So you don't want to focus just on the person, the single person. You really need to look for people in groups. Because people moved in groups. They didn't really strike out on their own necessarily. A lot of people ended up really not moving very far. — P5E

A third, more passive approach genealogists brought up was to wait until new evidence surfaces, as there are still many records to be digitized and published. Waiting may sound easy, yet genealogists expressed that they need to deal with negative emotions associated with not being able to solve a problem right away:

> It's hard sometimes. You just have to set them aside for a bit until more record collections come out. And sometimes things will appear, and then you can solve it. But sometimes you still never can. . . . It's frustrating, sometimes, though, when you can't solve things you know you should be able to. — P1E

Some amateurs participants admitted that they were driven by curiosity (e.g., "how far back can I get?"), for fun (e.g. "playing around"), seeking connections to historical significance (e.g., "am I gonna find somebody who was at one of the original colonies?"), or aching for an answer to their specific question about ancestor (e.g., "I was maybe too impatient to try to get the answers."). For them, the challenge is not to give in to negative feelings accompanied with not achieving their goal, **rush towards insufficient research, or simply give up**.

4.1.5 DNA as new evidence. Genetic testing has emerged in the genealogy community, and these services are now commercialized and integrated into genealogy websites. As people cannot lie about their genetic information, DNA testing is embraced by genealogists as a new type of evidence to complement historical records, which is especially important when there is little other information available, as in researching for orphans, adopted children, or those who born out of marriage.

In the meantime, historical records remain the major source for genealogy research. One expert emphasized that historical documents are necessary to explain what is presented in the DNA data:

> [DNA] is just another piece of evidence. It's not any more valuable than a birth certificate. You can get inferences from DNA, but you can't prove anything with just DNA. ... You've got to try to find some paper records that give you some insight into why the DNA comes out as it does. — P6E

Genealogists acknowledged that individuals' DNA should be compared with others', and the analysis could become more accurate as more people take DNA tests. As a result, genealogists understand that they might have to wait for more people taking the test to solve their research questions. Some even consider sending DNA test kits "strategically" to relatives as a gift and try to persuade them to take the test.

4.2 GPS Point 2: Complete and Accurate Source Citations

Keeping complete and accurate source citations for research requires good management of information during genealogy research. In practice, we find that genealogists need to document and organize their research and findings along the way for later reference and analysis.

4.2.1 Organizing with file systems and family trees. Organizing research findings was no easy task considering the amount of information genealogists need to collect. Participants mentioned that they store and organize digital records with local file systems and cloud storage. Family trees, either built with local genealogy software or on genealogy websites, naturally present an organized hierarchical structure, and genealogists can attach documents, citations, and add notes to ancestors' profiles in the tree.

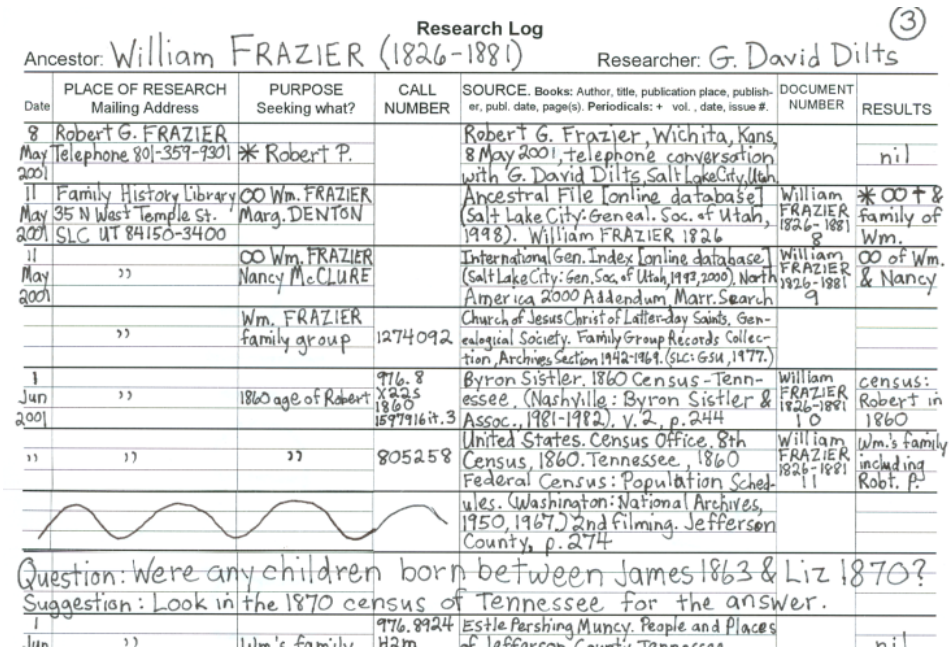
> During research, the Ancestry tree was actually a pretty active part of the process. It was where we ... stored all the information relating to people and families and stuff. If there was a document that we found if it was on Ancestry, we would save the Ancestry version to the Ancestry tree. If we found it elsewhere, we would either download the image and upload it, or we would scan it and upload it to the tree. — P8E

Organizing documents and citations can be tedious and time consuming. Many amateurs struggle at this task, as they often described their research being "messy," and one expert commented on amateurs that "half the battle is because they're not organized" (P1E). We observed that amateurs may treat organization as a secondary task, and **make little effort in organizing** their research. As one amateur noted:

> It's just like a really ugly draft of my research. And I don't want to put a lot of effort into making it pretty and "putting the bells and whistles" on it just to find out it's wrong, or I'm missing something, or I gotta kind of redo it. — P14A

However, as we will demonstrate in the following themes, experts also face the challenge regarding the uncertainties of their pending conclusions in their research, and they develop a set of tools to help accommodate it.

4.2.2 Keeping track with research logs. Many expert genealogists asserted that not only findings are worth documenting. In fact, they developed *research logs* (see an example in Figure 1), which are spreadsheets that document genealogists' searching activities to help them keep track of research progress, even if they end up with no significant findings. The benefit is multi-fold according to experts: First, the logs record information about their search attempts, such as what collection is searched, what keywords were used, in which websites, etc., so that they could avoid duplicating the search effort when picking it up later. Second, a research log states the current research question



Research Log (3)

Ancestor: William FRAZIER (1826-1881) Researcher: G. David Dilts

Date	PLACE OF RESEARCH Mailing Address	PURPOSE Seeking what?	CALL NUMBER	SOURCE. Books: Author, title, publication place, publisher, publ. date, page(s). Periodicals: + vol., date, issue #.	DOCUMENT NUMBER	RESULTS
8 May 2001	Robert G. FRAZIER Telephone 801-359-9301	✳ Robert P.		Robert G. Frazier, Wichita, Kans. 8 May 2001, telephone conversation with G. David Dilts, Salt Lake City, Utah.		nil
11 May 2001	Family History Library 35 N West Temple St. SLC UT 84150-3400	○○ Wm. FRAZIER Marg. DENTON		Ancestral File [online database] (Salt Lake City: Geneal. Soc. of Utah, 1998). William FRAZIER 1826	William FRAZIER 1826-1881 8	✳ ○○ † & family of Wm.
11 May 2001	〃	○○ Wm. FRAZIER Nancy McCLURE		International Gen. Index [online database] (Salt Lake City: Gen. Soc. of Utah, 1993, 2000). North America 2000 Addendum Marr. Search	William FRAZIER 1826-1881 9	○○ of Wm. & Nancy
	〃	Wm. FRAZIER family group	1274092	Church of Jesus Christ of Latter-day Saints, Genealogical Society. Family Group Records Collection, Archives Section 1942-1969. (SLC: GSU, 1977.)		
1 Jun 2001	〃	1860 age of Robert	976.8 X22s 1860 1597916 it.3	Byron Sistler. 1860 Census - Tennessee, (Nashville: Byron Sistler & Assoc., 1981-1982). v. 2, p. 244	William FRAZIER 1826-1881 10	census: Robert in 1860
〃	〃	〃	805258	United States. Census Office. 8th Census, 1860. Tennessee, 1860 Federal Census: Population Schedules. (Washington: National Archives, 1950, 1967.) 2nd filming. Jefferson County, p. 274	William FRAZIER 1826-1881 11	Wm.'s family including Robt. P.

Question: Were any children born between James 1863 & Liz 1870?
Suggestion: Look in the 1870 census of Tennessee for the answer.

1 Jun	〃	Wm.'s family	976.8924 H2m	Estle Pershing Muncy. People and Places of Jefferson County, Tennessee		nil

Fig. 1. An example of paper-based genealogy research log from [112]. Note that it contains the name of the ancestor being researched, a research question, and detailed documentation of means, purpose, and result of the search.

for this research session to help genealogists stay focused. Third, research logs help genealogists store information they found that is not directly related to the current question but might be useful as the research progresses, so that they can refer to it when necessary. Finally, for professionals, a research log helps them demonstrate to the client what they have spent time for during the research session. This is important as most professional genealogists charge based on the time they spent on the client's project. One expert shared with us a very detailed use of research logs:

> A lot of genealogists, professionally, use some kind of a log to track their searches. . . . And I just keep track of what I looked at, what source I consulted, and then I write down what was my purpose and intention for the search, what was I looking for. And then I also track what were the specific criteria that I used. So if it's a database, "Did I search for all people with just the first name?" "Did I look for all people with the last name?" "Did I use a wildcard to give me spelling variation?" So I'll keep track of the search parameters, just so that I know how thorough my search was, and then I'll make notes about the outcome. "Did I find some potential candidates?" "Did I find the ancestor I wanted?" Or maybe I didn't. So I keep track of all of that, and I make comments. — P9E

None of our amateur participants mentioned keeping research logs like experts did. They often seemed unaware of the benefit it may bring to them. Amateurs proposed alternative approaches to help them keep track of the research progress, such as appending notes and tags to family trees, writing down paper notes, and keeping browser tabs open. Unsurprisingly, they seemed to struggle at times and found this approach ineffective:

> I think to myself, "Oh, this is great. I need to remember to put this over for this person" and then. You get distracted, you have to go do something else, something comes up, then you come back the next day and you go, "What was it I was thinking about?" So I don't currently keep a log or a record of what I'm doing and I should. — P19A

> I'm the person that'll have 30 tabs open, and I'll be opening tabs until it forces me to stop, or, you know, I've got a little notebook next to me that I'm making weird gibberish notes. And then if I go back a week later, I'm like "I don't know what the hell this means!" — P15A

4.3 GPS Point 3: Thorough Analysis and Correlation

The third point of GPS suggested that genealogists need to evaluate (through analysis) and summarize (i.e., correlate) all the information collected that is relevant to the research question to draw a conclusion. Expert genealogists take cautious steps in this process, noting that research of each ancestor is built upon that of closer generations in the family tree. Genealogy literature introduced metrics in evaluating the *source*, *information*, and *evidence*, each of which has a specific meaning in experts' genealogy research. As a primer, sources are records in various forms which could be *original* if the record is not derived from other sources, *derivative* if it is, or *authored* if the author's opinion and observation is involved. A piece of information is considered *primary* if the informant has firsthand knowledge of the described event, *secondary* if they does not, or *indeterminable*. Lastly, evidence is the body of records that are relevant to the research question, which can be categorized as *direct*, *indirect*, or *negative* based on whether the evidence could directly answer the question [91].

4.3.1 Knowledge of information production. Expert genealogists demonstrated awareness of how information was produced and how errors and mistakes could be made in the process throughout our interview study. They described in detail the cultural background of historical records and why the information might not be reliable in some of them, the process of original documents being digitized, transcribed and indexed for easy access and search, and how record hints are generated from indexed records and others' trees on genealogy websites. Genealogists essentially use such knowledge to assess the quality of information they gathered. The following examples shows how deeply genealogists understand the information they are working on. For instance, on original documents:

> "…Well, it's in the census record. It has to be right!" And I said, "Look at that low X shows who gave it. It was an 8-year-old daughter. 'Tell me when, [in] what year, your parents were born in Ireland!' " — P2E

On record hints:

> "As far as record hints go, that is only as good as the computer algorithms and what is already in the tree. …So it's only searching through the small amount of indexed information, and matching it to what people have put in the tree, whether it's right or wrong." — P3E

Amateur participants did not describe the ins and outs of the information production procedure as precisely as experts. It appeared that amateurs **sense the reliability of documents from experience**, who have also expressed their uncertainly in assessing some of the records they encountered:

> I have found they [church records] are very thorough in terms of noting the names of everybody involved, birth dates, christening dates and so on, marriage dates. They're really pretty thorough at that time. So I do think they're accurate. But then there's

> other documents, like in 1920 [census] in the United States, that I'm not quite sure are as accurate. — P20A

4.3.2 Verifying clues by oneself. Even though expert genealogists are willing to leverage research hints and others' trees in their research, they are also selective when evaluating if one piece of information can serve as evidence in formulating their conclusion. In general, experts expressed their preference of using primary information from original records as evidence, while considering the rest, such as others' family trees and algorithmic research hints, as research clues that they need to verify themselves. In P6E's words, "I'll take it [someone's tree] as a hint. And I'll do what I have to do to validate it."

Besides re-approving everything in others' family trees, genealogists also develop strategies to effectively assess the overall quality of a family tree. In particular, they mentioned several defining characteristics of a tree, such as sparse supporting evidence, conflicting or duplicate documents in a profile, and overly large or deep family trees, indicating that the tree might be poorly researched or copied from somewhere else. For instance, one participant described using the ratio of records to ancestors in the tree in his evaluation metrics:

> There's no way most people have enough records, enough proof of who 6,000 people are to put them in there. . . . it'll show, say, 6,000 people, 6,500 records. So you go, "You have barely over one record per person. 'Cause I have people with 20's and 30's [records per person], that tells me that you have not clearly looked at their stuff." — P19A

Additionally, experts asserted that amateur genealogists who are less competent often **treat clues as evidence without proper verification**. Several amateur participants of our study admitted the problem of they accepting research hints to build family trees without any verification at the beginning of their research. P16A reflected on his initial research on genealogy websites:

> I did initially, but I no longer sort of blindly accept the connections that other users on Ancestry did. Like the first weekend that I had my Ancestry account, . . . you have like a little shaky leaf that creates hints, and I thought that I had successfully traced my ancestry back to like 1200 [CE] England. — P16A

Participants further acknowledged that such behavior is common among amateurs. As P18A elaborated, "There's a big problem in genealogy community at most of the beginner's stage . . . where I'll just copy information over and over." Considering how easy for users to accept research hints and duplicate others' family trees on genealogy websites, one can quickly expand his/her family tree by replicating unverified information. However, expert genealogists emphasized that the frequency of one piece of information does not imply its accuracy. P1E argued, "Just because it's there 40 times doesn't mean it's true."

4.4 GPS Point 4: Resolution of Conflicting Evidence

Even when working with original documents, genealogists will likely find conflicting information among the records. Participants mentioned two types of conflicts. First, they suggested that various records may not belong to the same research subject, even though partial information (name, age, location, etc.) matches. Second, there are also cases where conflicting information exists due to the inaccuracy of historical documents, as discussed in the previous theme. Therefore, it relies on genealogists' expertise to determine if all the information they collect is accurate and actually belongs to the same person. Furthermore, circling back to the first point of GPS, only after reasonably exhaustive research can genealogists confidently determine that all conflicting information is identified and resolved before reaching the final conclusion.

4.4.1 Keeping an open mind dealing with uncertainties and new information. First and foremost, experts acknowledged that following GPS does not guarantee perfect conclusions, as new records might appear at later time that overturn a conclusion:

> The genealogical proof standard doesn't guarantee that your answer is correct. But it makes it more likely that it is correct. It's more likely that if new information surfaces that your conclusion will stand. — P7E

As a result, expert genealogists try to keep their mind open to other interpretations and new pieces of information. One participant stated:

> A lot of genealogists don't ever say they've proven something. They say "the evidence shows," or "it looks like this is the answer" or "this is probably the correct thing." And we're always open to more interpretation in the future as more records become available. — P3E

Additionally, several participants reported that they often met amateurs who **hold predetermined beliefs** from family stories and their own research, which prevented them from accepting evidence that indicated otherwise:

> When people hear a family story sometimes they have a really hard time letting go. . . . If people get so fixated on, you know, these things that they've heard, and they've believed to be true, that sometimes, even when you walk them through how records show that it would be impossible, they just don't believe you. — P4E

Year	Month	Day	Who	Event	Location	Notes	Source
1870			Terrence Lee	residence	Deerfield, MA	44. Honora 42. Patrick 11, Ellen 11, Terrence 4, Bernard 7, James 20	1870 census
1870			Margaret Lee	residence	Deerfield, MA	Francis 60, Margaret 58, Francis 26	1870 census
1870			Catherine Lee	residence	Deerfield, MA	Kate 23, william 27	1870 census
1871			Philip Lawler	marriage	Deerfield, MA	24, born in Ireland	marriage register
1873	Nov	20	Terrence Lee Jr	death	Deerfield, MA	9 years old. Buried Green River. Diptheria. No location for births. Not on Findagrave	death register
1874			Patrick Lee	Naturalization	Unknown	Massachusetts or Colorado?	1920 census
1874	Nov	24	Patrick Lee	marriage	Northampton, MA	Patrick 22 living in Deerfield, born Ireland. Alice 21 born Ireland. Married by ME Barry priest. Marriage register. Find church records	marriage register
1876			James S Lee	marriage	Holyoke, Hampden, MA	30 Nov. James 26, Mary 25. Both born Ireland. Married by Rev PJ Harkins of Holyoke. Find church records.	marriage register
1876	August	14	Terrence Lee	death	Deerfield, MA	53, born Ireland. Parents James and Mary Lee. Buried Calvary. Find picture of grave	death certificate
1880			Bernard Lee	residence	Holyoke, Hampden, MA	18, born Ireland	1880 census
1880			Patrick Lee	residence	Leicester, Worcester, MA	26, born Ireland. Alice 26, born Ireland. James 4, Mary L 2	1880 census
1880			Catherine Lawler	residence	Greenfield, MA	34, born Ireland	1880 census
1880			Mary Ann Lawler	residence	Peabody, Essex, MA	42, born Ireland	1880 census
1880			Honora Soraghan	residence	Holyoke, Hampden, MA	55, born Ireland	1880 census
						Catherine 33, william 36, both born Ireland.	

Fig. 2. An example of genealogy timeline from [80]. It lays out chronological and geographical information of multiple ancestors to help a genealogist to identify research gaps and conflicts.

4.4.2 Resolving conflicts with genealogy timelines. In practice, genealogists developed the timeline tool, which is a spreadsheet that lays out ancestors' lifetime activities in sequence based on the collection of records (see Figure 2). By presenting information in this way, expert genealogists can spot conflicts by identifying events that are impossible to occur given natural constraints and historical context, such as children being born before their parents or people moving between distant places within a very short period of time. Furthermore, the timeline tool can help genealogists to recognize missing information in certain periods of an ancestor's life. As one expert summarized:

> Timelines are another really important tool that we use. It's particularly [useful] if you have two people with the same name when you're trying to separate them out and figure out which John Smith is this one. . . . If you lay out a timeline, and where they are, and what the date is, you can sort of figure out which of these facts goes with this particular person. . . . It also helps you figure out where your holes are. — P5E

Similar to research logs, amateurs **didn't specify creating timeline sheets themselves**. We notice that some genealogy websites provide timeline functionality for their users. However, these timeline features tend to be constrained to an individual ancestor's level, which is not as flexible as experts' customized spreadsheets, and may hinder genealogists' analysis across multiple research subjects.

4.4.3 Testing hypotheses with family trees. When exploring alternative explanations, several genealogists also mentioned that they would build hypothetical (or speculative) family trees to help them test their research hypotheses. We heard different approaches from our participants on this matter across our interviews. One participant suggested that she would use the ancestor profiles to store unexplored leads:

> Once I get the first document that indicates, "Okay, they might be related to me through this other person," then I will create a profile for them, basically just adding their name to the tree and who they're connected with, but there really isn't much on it. I don't add in any facts until I can find the documentation for them. So I just mainly create the little profile just to keep track of their name. But no, I don't really rely on it until I can find more information. — P20A

Another participant mentioned that she would build the ancestor profiles first by accepting hints, and resolve conflicts in it afterwards:

> I would kind of do a rough draft, accepting hints, and then kind of go back through it and check my work and add legitimate records and census records. If I couldn't find anything then I would delete it. But, usually, ... if you come back a couple of days later, there's gonna be more hints for records, and then you can kind of cross reference your records, and make sure that everything fits. — P15A

We note that **hypothetical work is often part of users' public family trees**. In such cases, other genealogists may have to refer to notes and tags in the tree, or rely on their own interpretation to determine which part of the tree is thoroughly researched and which is not. Some genealogy websites allow building multiple trees and keeping them private, so users have an option to detach hypothetical work from the public (i.e., proven) family tree. And one expert mentioned working with two genealogy software to separate proofs and hypotheses.

4.5 GPS Point 5: Soundly Written Conclusion

The final step of genealogists' research is to express their research conclusions in appropriate formats. Expert genealogists treat their research product seriously, tying the quality of it to their professional reputation.

4.5.1 Producing research reports and family tree data. Research reports are a very common product in expert genealogists' work. The format of a report may vary based on clients' needs, but generally, it will contain all the (references to) documents genealogists have collected that are related to the research question and a thorough explanation — expert genealogists called it proof argument — on how they reached the conclusion based on the evidence. Expert genealogists may publish their research (sometimes on behalf of clients) in genealogical journals or as books. They use research reports to showcase how they follow the GPS throughout the research, and others can also use it to evaluate the quality of work:

> You have to give people enough information, preferably, so that they could go and find it [evidence] themselves and confirm that you actually interpreted it correctly ... You have to give people enough information so that they can judge the quality of your

> work. If you see a genealogical work product without documentation, as is often said, that could be fiction somebody have just made up. — P7E

In practice, family trees and associated data sets can be an important complement to genealogists' research deliverables. Genealogists have developed open-source data formats, such as GEDCOM (an acronym for Genealogical Data Communication) [42, 113], that are widely accepted by the community for storing and transferring genealogical data. In addition, an expert noted that online family trees have made it convenient for them to keep clients posted about the research progress.

We found that many amateurs are **enthusiastic in sharing their findings online, but not necessarily in formal reports**. Their sharing is motivated by the intent of reaching out to relatives and altruism to descendants and other genealogists, as one participant mentioned:

> I really wanna share what I find. I know it's maybe a small group of people who might be interested in that, or someone might be interested in it years from now. . . . Because it's no harder for me to do that than it would be to keep my own records and then it has the added benefit of maybe helping someone else now or down the road. And certainly I've appreciated that when people have been transparent and had public trees and things like that on Ancestry or on FamilySearch. — P16A

Similar to organizing, experts agreed that producing a report with full industry-standard citations is "very time-consuming", and some of them may leverage genealogy software to automate part of the report generation process. Amateurs, on the other hand, simplify the process by self-publishing family stories or appending records and notes to their family trees to express their conclusion. Meanwhile, our findings suggest that there might not be a community norm on how a family tree should be built and presented on the Internet. It could be a combination of verified work, hypotheses, and research leads, depending on individual genealogist's preference.

4.6 Community Interactions

Social context is important when investigating the learning experience of community members. In genealogy specifically, as one participant commented, "there is a huge social component," noting that the self-identity aspect of genealogical research is "super tender to people" (P11A). In this section, we focus on social interactions in this community, rooted in the practices of genealogy research that surfaced throughout the interview study.

4.6.1 Privacy concerns in online collaboration. It is common for genealogists to seek help from their peers in online groups. Our participants shared with us several reasons: specialized or technical assistance such as analyzing complex records or DNA data, translating documents from a foreign language, and restoring aged photos; as well as seeking research advice more generally. Many of these services (except for image restoration) are provided by community members free of charge. Similar to genealogy websites, amateurs may decide to jump across different online groups based on their specific information needs. Genealogists expressed great appreciation of these online groups for their altruistic work:

> I'll send my students there [a Facebook group] because it's really really helpful. If I have a question, people can answer it, and there's experts there that will answer and they do it very quickly. So it's a great way to get help, and with little cost or no cost. — P1E

We heard about diverse practices for sharing genealogy information from our participants. For some amateurs, their enthusiasm in sharing genealogical information goes beyond genealogy-focused websites, extending to general purpose websites like social media. Several participants expressed concern for the **risk of revealing personal information** in genealogy research, who

may limit their appearance on social media and provide just enough information to an online community when asking for help. It may make the online collaboration less effective, but their privacy concerns are not baseless. Government records are usually released to the public decades after they are generated for privacy reasons [44]. Genealogy websites follows a similar rule, limiting access to profiles on living people in the tree [1, 2]. In the meantime, more than one amateur participant mentioned researching living individuals who they think might be their relatives using public online information:

> I then, to be honest, I then get online, get on Google. If they have any information at all, you know, if they have a unique name — I've also tested at 23andMe and so if they're on 23andMe then there's usually a little bit more information — sometimes there's their age, sometimes it's what city they live in or what state they live in and I can just start Googling their name and a potential age and start to identify them. ...I'm basically, if you are familiar with the term, "doxing" someone. — P17A

4.6.2 Peer review and conflicts. Some expert genealogists emphasized the importance of sharing and peer-review, in a scholarly or publishing context, as a means to ensure genealogy research quality. Some experts expected this approach to be applicable on genealogy websites as well, while one expert acknowledged that the growing diversity of sharing mechanisms makes community moderation challenging:

> The way that genealogical information is shared has really moved towards these forums, websites, or self-published materials. You're not really seeing it mediated quite so much by genealogical society. — P8E

In practice, many genealogists expressed **great difficulty in correcting others' (potential) mistakes** on the Internet. First, it is almost impossible to correct inaccurate information propagated across data sets and individuals' family trees online. One expert shared us an example of himself:

> That piece of information has been copied through GEDCOM and GEDCOM and GEDCOM, and copied up into Ancestry file after Ancestry file. Everyone's pulled it up on FamilySearch. ...They've copied it at MyHeritage. It's just out there and there is no way to pull that back. ...It's kind of like standing in front of a landslide and going, "Stop! Stop!" It isn't gonna happen. — P5E

On the other hand, recipient of critics and correction suggestions could feel offended and ignore the message. For instance, a participant expressed frustration when sharing her most recent findings on r/Genealogy subreddit, for other people suspected there might be mistakes in the conclusion:

> I just don't know if I wanna go back to justify to that group how I know what I know. ...Some of the posts about it were not very polite ...I'm happy to maybe answer questions if you didn't try to insult me about how I found this so quickly. ...Even as helpful as those people are on other things, every once in a while they ...just get a little rude, a little toxic. — P20A

Last but not least, many genealogists who found potential mistakes often feel hesitant to contact the author, in fear of hurting others' feelings, messing up another user's tree, or getting involved in an argument, regardless whether they know the author personally:

> Within the past weekend I found a huge error on one of my cousin's trees, and I haven't been able to tell him yet. I don't know how, and I know he's spent years and years on this research, but he got one of the people wrong in the very beginning. ...And with this particular cousin I have editorial status that he's given me on his tree. So I can actually go in and fix it, but I don't wanna mess up all of this work. — P11A

> I'm not confrontational, so I'm not gonna call anyone out. . . . On one hand, I probably take it more seriously than I should, because I just get obsessed with [genealogy]. But, on the other hand, I don't care enough to get into an Internet fight with anybody about something. — P15A

4.6.3 Perceptions of expertise. Expert genealogists are not free from the conflicts described in the previous theme, even when working with or for amateurs. Experts shared with us their experience when clients were not satisfied or disagreed with their research outcome. They did not find solutions to resolve the conflict, except for acknowledging different opinions or being more selective taking clients. In one case, P4E recalled, "My client didn't believe me, and so we very graciously and politely decided to part ways." P9E reflected, "Through my years of experience, I'm a lot better at selecting the clients that I wanna work with, and that I know will be good partners and working with me."

Similarly, genealogists' expertise does not grant them special authority in most online environments. They understand that the family trees they build on open genealogy sites are modifiable by anybody else, regardless of their research capability. One expert also stressed the challenge of content moderation on genealogy websites, and hoped that effective education of the amateurs might help mitigate the problem: "There's not a really good way to say 'you're banned from posting about genealogy for the rest of your life,' but I feel like the more people I teach how to do good research the better the research will be eventually" (P3E).

A few experts, however, intended to **disengage from intensive online interactions** with the public — often after an initial attempt of sharing their knowledge — to avoid the wariness and complications. As one participant summarized:

> There are genuinely kind people that will donate their time and share this advice freely. And there are plenty of people like that. But it seems to me [that] people with the most experience tend to veer towards really just providing their analysis while they're on the clock, . . . or operating as a specific capacity for their job . . . I think a lot of us professional genealogists all just kind of go through this initial phase where we join every group possible and want to share everything we've learned, and then we get burned out. And then we just focus on our work in our own little corner. — P8E

We also asked participants about financial relationships between genealogists. We found that not every amateur would like to hire experts to research for them. Cost is certainly one factor of concern. More importantly, they expressed **preferences of researching by themselves** to enjoy the experience of solving mysteries, gaining knowledge, and exploring historical context that connects to ancestors' life through genealogical research:

> The fun of it is discovering it for myself instead of just having someone hand it over for me. It's just being able to feel connected and immersed in, and then feeling good and enthusiastic about the connections or things that I've been able to find or validate ... So I wouldn't be opposed to it [hiring professionals]. I think that they bring a ton of value, but I definitely think it takes away a little bit from the experience. — P14A

Those who consider hiring experts may do so for acquiring local documents, overcoming brick walls or double-checking their own research. Amateurs don't blindly accept experts solely based on credentials. Instead, they **assess experts' research capability** using various criteria, such as how physically close experts are to the record collection, the professionalism of service's website, and affiliations with prominent organizations.

4.7 Genealogy Learning and Education

As we have discussed, amateur genealogists often conduct original research on family history, and individuals' research ability has great impact on the quality of the overall product on genealogy websites. Therefore, understanding how amateurs learn to conduct good genealogy research in a technology-mediated context is critical for this community.

4.7.1 Learning by doing. Learning-by-doing is a common approach mentioned by both amateurs and experts, many of the latter having started genealogy research as a hobby. Amateurs described "jumping into" genealogy without much knowledge about proper research methods, relying on what the software or websites provided them. In P15A's words, "Honestly, I guess Ancestry.com itself is … just user-friendly enough that I figured it out."

Amateurs admitted that they could make **lots of mistakes at the early phase of research** with this trial-and-error approach. Their learning is triggered when realizing previous mistakes or encountering research barriers. As their knowledge grows, amateurs described researching more rigorously to review, improve, or completely redo previous work. However, we also found that some of the early mistakes could remain in the family tree. It was unclear whether every amateur would recognize their mistakes and start learning the proper research method soon after they started genealogy research.

4.7.2 Seeking the right learning resources. Genealogists may need to learn not only research methods, but also facets of many other subjects, such as foreign languages and history, to advance their research. Amateurs' learning is typically self-directed. They need to **proactively look for content that is appropriate to their knowledge level and information needs**. For example, P13A said, "I don't know where my German ancestors came from, so I need to learn about the history of Germany, so I can narrow it down by what region they might have come from."

Amateur participants mentioned free online resources for learning genealogy research, like YouTube videos, podcasts, and blogs, with video lectures emerging as among the most common sources for learning. Even if amateurs did not actively participate in online groups, they considered browsing others' posts an important and effective method to learn from them:

> I look for posts … and read the follow-on posts for further ideas about how I can do the work that I'm interested in doing. I look at it, I monitor it and I read the comments looking for more ways to to expand my research to do something better. — P17A

Experts endorsed many of the free online resources that amateurs mentioned. Furthermore, a few of them proposed that the FamilySearch Research Wiki [111], a web-based encyclopedia of genealogy research, could be a good source for learning, but it was rarely brought up by amateurs when they described their learning. One amateur, P13A, even explicitly requested a "Wikipedia style" site apparently without the knowledge of FamilySearch Research Wiki, which may indicate a disconnection between learners and learning resources.

Experts also suggested more formal resources for newcomers to advance their research, including seminars offered by genealogical societies, degree and certificate programs from higher educational institutions, genealogy journals, and professional study groups. Depending on amateurs' allowance and research progress, formal education might become more of an option when their knowledge on and dedication to genealogy increases. P20A explained her evolving priorities:

> I think I'm probably reaching the point where I would go to seminars and things. I was maybe too impatient to try to get the answers that I wanted at the beginning. And now that I've got a lot of those answers. And the urgency isn't there quite so much now, I might be willing to go and just learn about the technique of genealogy research.

4.7.3 Knowing without implementing. We also heard from amateurs that they may **not apply their knowledge on genealogy research in practice**, which suggests that having knowledge on genealogy research methods does not directly translate to their implementation. As P16A mentioned, "I've listened to some genealogy podcasts and I've seen some webinars. I'm familiar with the formal research process, . . . I have never gone about researching genealogy in such a formal way, like by having a literal question that's written down with supporting evidence."

One reason we identified is that amateurs may under-appreciate the benefits of established research methods due to less experience and a misalignment of research goal and method. For example, one participant spoke to the tension between her primary motivation to tell a good story about the family, and the suggested method of explicit and structured organizing and analyzing:

> There's so much out there and I've watched a lot of YouTube videos and tutorials on the tried and true and the best practices and how to organize and lay it out. I struggled with that method a lot in the beginning because it wasn't my vision. A lot of it . . . have these very structured trees, and it's just the facts, and that just goes down the lineage. But to me, I didn't connect with that. I wanted my book to tell a story about their lives and their experiences, and how it related to events, and really be able to connect. — P14A

4.7.4 Learning with genealogy software systems. A couple of expert genealogists asserted that amateurs can accomplish much more nowadays. Optimistically, they suggested that modern amateurs' skills can sometimes be on par with experts of previous eras, thanks to emerging technologies in the genealogy industry:

> So many people now can get started and make a lot more progress that 20 years ago, I know professionals who were working 20 years ago that aren't working now, because the clients have matched or outpaced their skills. . . . The client can now come to us with so much more prep work than they used to, because now they know how to, where to go, and they can get so much more very quickly. — P9E

In contrast, amateurs themselves asserted that **genealogy software systems could have been more helpful** in educating newcomers about genealogy research.

> I wish Ancestry did a better job of explaining basic methods as you're starting a tree . . . They haven't educated the consumers on what to do before starting this. . . . They're just like, "Here, go at it!" without kind of explaining the methodologies, and why you cross reference all these different things to make sure you have the same person. — P11A

> I got DNA tested on ancestry. I got DNA tested on 23andMe. And I don't believe either one of them had a good "Here's how to get started." . . . I don't believe either of those sites have that, from a beginner point of view, "what to do" and so forth. — P17A

We asked participants what functionalities they would like to see to help with amateurs' research. From a high-level view, experts believe that amateurs would benefit from "hand-holding" instructions with prompt question-answering during genealogy research:

> A lot of beginners are just like, "Well, I just want a step-by-step thing, maybe like with the video, or just where I can stop and ask questions." And I think that's something where technology would play a huge role in automating that and then updating that. — P4E

Amateurs provided more detailed requests. They imagined that AI-powered computer systems could explain proven research approaches to users, propose research directions, introduce new sources, and present historical context during the research:

> Tying it to historical events definitely brings a lot of value. And so if you're pulling these census or pulling these certificates, or you know these immigration record or whatever, . . . the AI would be "hey, these are also recent historical events." And I know Ancestry does have kind of a similar timeline, but I don't think it's integrated when you're searching. — P14A

> [If] I buy an old tax record, I can't be sure if it's accurate, right? Maybe AI could help me determine if it is accurate, and then beyond that could tell me what next I could do with it, and give me 3 reasons why that's relevant to my life today 150 years later. — P16A

Automatically verification of records and timeline is another desired features among genealogists. They suggested that computer systems could flag questionable conclusions and conflicting evidence to help amateurs improve their research quality and assess the quality of other's work more effectively. In P10E's words, "They could have a verifiable system where you would have some reliability, and that would also help the amateur, because it would force the amateur to go do the research."

5 Discussion

Wenger [110] described Communities of Practice as "formed by people who engage in a process of collective learning in a shared domain of human endeavor." As genealogy research practitioners build family histories with available information from various types of historical documents, we describe the genealogist community as a *community of information practice*, in which genealogists learn best practices involving information foraging, analysis and production. Our study revealed that genealogy research practices resonate with frameworks of sensemaking and information practice proposed by prior research in other domains. For example, Krizan [62] proposed an intelligence analysis process that comprises the conversion of information needs into intelligence requirements, in addition to information collection, processing, and production. Pirolli and Card [87], in their influential work on sensemaking, also described a process of foraging information, extracting evidences, developing structured schema and hypotheses, and constructing final presentation. From an information literacy perspective, the Association of College and Research Libraries (ACRL) also stressed the importance of understanding information creators' expertise and credibility, the necessity of managing information, and awareness on ethical and legal use of information [83, 84]. All of these elements were reflected in genealogists' research routines in our study.

Yet, the challenges our participants experience are not exclusive to the genealogy domain. Amateur genealogists' approach to accepting hints echoes Kang et al. [50]'s observations of "teammate inaccuracy blindness," which in turn leads to the propagation of inaccurate information in an asynchronous collective sensemaking process. Li et al. [66] also discussed an increasing number of errors in their crowdsourced sensemaking pipeline due to "insufficient context and missing information" from early steps. Their findings on different types of errors, such as inattention to background knowledge and low effort errors, also find their counterparts in amateurs' genealogy research behaviors. As amateur genealogists struggle with assessing quality of historical documents due to their insufficient knowledge, Koesten et al. [60] similarly stressed the importance of data creators communicating the original purpose and limitations of data with its users. Our study contributes evidence towards the universality of these collective sensemaking challenges at an online community scale.

Importantly, experts in the genealogist community developed research standards heuristically, attempting to ensure quality of the community's research outcomes by regulating essential aspects of information use and production for individual genealogists. Experts advocate for explicit distinction

Category	Challenges of Amateur Genealogists
Inadequate research	• Unfamiliarity with records and research methods • Broad, shifting research questions • Scarce, unorganized documentation of research process and conclusion • Insufficient analysis and correlation of evidence • Impatience and biases • High-productivity, low-accuracy research
Barriers of community interaction	• Unestablished norms on how to use family trees • Conflicts between genealogists • Recognizing and acknowledging expertise • Privacy concerns
Challenges in learning	• Locating learning resource that fits one's knowledge and needs • Adopting appropriate research methods • Leveraging genealogy software systems

Table 3. Summary of challenges faced by amateur genealogists to research effectively and conduct high quality research.

among different types of information sources (original documents, research hints, others' family trees, etc.), critical evaluation of information (based on its quality, correlation with other information, and relevance to the research question), and thorough documentation of research process and reasoning. Their standards and practices provide useful guidelines for us to identify challenges in amateurs' approaches and ways to support them in genealogy research. We summarized these challenges in Table 3 and explore mitigating approaches.

5.1 Technological Support for Research Quality

Our findings on genealogy research are largely consistent with previous studies, including expert genealogists customizing research tools [23], searching for local resources [120], the requirement of history knowledge [32], and the importance of information organization and production [34]. We also confirmed that challenges resulting from inadequate genealogy research reported by Willever-Farr and Forte [116] persists a decade later, despite the emerging DNA testing and increasing availability of educational resources on the Internet. While our interviewees with professional backgrounds of lawyers and government analysts asserted that the "analytical skills" they acquired from previous careers are related to genealogical research, we also find common challenges with other fields of study, such as asymmetrical skepticism (confirmation bias) [72, 94], emotional proximity (connection with self-identity) [47], and digilantism (unsanctioned investigations of others) [81, 101].

While genealogists praised the conveniences afforded by genealogy websites, we also found negative consequences of it. Specifically, we identified a high-productivity, low-accuracy information production phase at the very beginning of amateurs' research, in which complete beginners extensively utilize research hints provided by algorithms and build large family trees instantly. In this case, inexperienced genealogists and less reliable research hints amplify one other to exacerbate the situation. Unfortunately, such misinformation is extremely difficult to retract once is propagated. Even when amateurs are aware of their mistakes later, they may still keep these erroneous branches for future fixes, which may further hinder the effectiveness of the algorithm and mislead more users in a vicious cycle. We therefore encourage the development of recommendation algorithms

that are sensitive to users' expertise and research quality which could suggest more reliable sources for amateurs.

Analyzing evidence is another task that amateurs struggle with. Genealogy websites may educate their users regarding the production of genealogical information, such as transcription, indexing, and algorithmic suggestions, to assist the evaluation of information. Besides bringing contextual information, we suggest that the platforms may also provide a canvas for cross-referencing multiple evidences and ancestors for correlating evidence and resolving conflict information, which is currently accomplished with customized spreadsheets by expert genealogists, but absent for amateurs. Automatic verification is also expected to help genealogists identify conflict information and avoid mistakes more effectively. Furthermore, since building hypothetical family trees is common for genealogists to explore other possibilities, more explicit differentiation of experimental trees with finished work might be introduced with a refined user interface.

Genealogists conduct original research as experts emphasize self-verifying on all clues, including system suggestions, family stories, and others' work. As a result, we observe that the evaluation of genealogical research is not solely based on the conclusion, but also the process used to reach it. Thorough documentation and organizing of the research is necessary for this goal, and experts also argue that it benefits genealogists as well by helping regulate their research. However, these tasks are time-consuming and often treated as secondary objectives to amateur genealogists, possibly due to the high cognitive demand of exhaustive information search and analysis. Therefore, genealogy platforms may consider automating these processes for users to reduce their overhead, but still prompt information such as current research goal to help regulate amateurs' research.

5.2 Leveraging the Power of Community

Our findings suggest that the online genealogist community occupies two types of platforms simultaneously. The first is genealogy websites, where genealogists build and organize family trees and directly collaborate with other users. The other is social media and online forums, in which genealogists find like-minded people, seek help, and learn from each other.

Genealogists may interact differently on these two types of platforms. On genealogy websites, amateurs are enthusiastic in sharing their family information with other genealogists, who are potential relatives and collaborators. On social media and open forums, users' privacy concerns may restrain sharing the same information with the public. Such differentiation also suggests that both types of platforms have their unique strengths. Genealogy websites can support collaborative problem solving, should such features be integrated into the system. For instance, a user could easily share a segment of their family tree and tag the ancestor with whom they face research challenges, and others who are familiar with the family, time period, and geographic location could step in with a more comprehensive view of the inquirer's research state, identify undiscovered mistakes, and provide advice accordingly. Online forums, on the other hand, are well-suited for socializing and asking more abstract questions with less risk of exposing personal information.

Implicit and explicit social interactions [93] may help genealogists establish community norms and avoid conflicts, but are often absent on genealogy websites. For example, featured family trees and ancestor profiles can be shown to amateur users to help them understand the expected research standard and how they should use family trees in their research. Expert genealogists may also benefit from this approach, as they may use featured family trees and profiles to showcase their research capability to amateurs, thus establishing community consensus on research expertise. To support more explicit social interaction, peer-review features such as rating and flagging of family trees would allow users to assess each others' research product and provide feedback more conveniently, which may also offset some of the challenges in direct communication. Finally, as previous studies [4, 57] suggested that chatbots could be effective for moderating online discussions

and debates, similar technology could be deployed as a moderator to ease tensions and scaffold productive, evidence-based communication between genealogists who have different opinions, facilitating a more collaborative online environment.

5.3 Computer Supported Andragogy

Individuals' genealogy research often starts as a hobby, and we found amateurs' learning experience fits in well with the concept of andragogy, i.e., adult education and learning, as opposed to pedagogy, the education of dependent personalities such as children and students [29, 58]. Andragogy is characterized with problem-oriented, just-in-time learning, and leverages experience and knowledge of learners [29]. Indeed, we found that many amateurs prefer learning during research. Their learning is self-directed, motivated by the challenge they are facing, and they need to identify the right learning resource that is suitable for their particular knowledge level.

We highlight a few characteristics of genealogy research and the genealogist community that make educating community members critical and challenging: Genealogy research is *original* that every genealogist is going through the whole research process and expected to verify others' work before accepting it. It is *personal* that family history is often associated with family stories and self-identity that not to be easily altered by others. The community is *inclusive* that anybody, regardless of their genealogy research experience and literacy skills, is entitled to explore and publish their family history. Consequently, the overall quality of the collective product on genealogy websites is heavily dependent on the research capability of every individual. When the popularity of genealogy surged among millions of hobbyists, experts in this field faces significant challenges scaling up their traditional educational approaches.

Meanwhile, amateurs are not guaranteed to be exposed to the established research standards and resources with their self-directed learning. Since amateurs rely on genealogy software systems heavily at the beginning of their research, we believe that these systems play a vital role in not only quality control, but also amateurs' education. On one hand, they may enforce aspects of GPS to regulate amateurs' research practice. More importantly, it is critical to provide appropriate guidance and explanation throughout amateurs' research to help them aware more resources, understand the benefit of each research approach, and scaffold their learning of genealogy research. Machine learning based tutoring systems [99] and large language models (LLM) [51, 63] could provide more scalable and user-friendly communication interface for amateur genealogists, while their effectiveness on genealogy research education needs further investigation.

As genealogists gain awareness of their knowledge gaps and seek to advance their research, the coming challenge is to identify the educational content that fits their need and knowledge level. Learnersourcing, an approach designed to support collective knowledge sharing among peer learners while maintaining their learning experiences [36, 55], might be helpful for not only sharing insights on historical documents, but also recommending and synthesizing educational content among community members. Recognizing individuals' research experience in online forums could be challenging [115], while prior studies [77, 86, 121] suggesting that system logs may help characterize users' research behavior in classroom settings, which might also be applicable to genealogy software systems. Future studies could be developed to evaluate users' genealogy expertise and challenges, and provide proper educational content accordingly.

6 Limitations

Identifying less experienced genealogists in an online community is not as straightforward as it sounds. We recruited participants who do not have credentials from the most accessible online community we can identify, expecting to reach varieties of genealogists in the wild. We acknowledge that novices who have not joined any community are not reachable with this recruiting approach,

but we believe our participants' recall of their early research practices complemented some of the absent perspectives.

Recruiting from other communities and groups may introduce varying observations. However, our participants did not join Reddit or professional associations exclusively. They also shared activities in other online communities and groups, including Facebook groups, Discord, online genealogy study groups, local genealogy societies, and direct messaging through genealogy websites. Therefore, we believe our participants would be a valid representation of the broader online genealogist community.

Our study participants were all located in the United States at the time of their interviews and a majority of them are of white ethnicity. They might present a more U.S.-focused perspectives of family history research. Further studies could be conducted with genealogists from different geographical, cultural, and ethnic groups to expand our understanding of genealogy and family history studies worldwide.

7 Conclusion

In this paper, we conducted an interview study to investigate genealogists' research process in detail, explored how amateurs' research might differ from established standards, and how it may introduce problems in an online, collaborative environment. We then explored social context of genealogist community, experts' and amateurs' perspectives on genealogy education, and their expectations for genealogy software systems. We argued for the importance of technological support for quality control and educating newcomers of this community, and made design recommendations accordingly. As genealogists' independent research is aggregated by online genealogy platforms, this field offers a unique bridge between well-organized communities of practice and individuals' everyday information activity. More preciously, the genealogist community emphasizes explicit research standards and thorough documentation of research process. We believe that this community will provide HCI researchers a new lens through which to examine challenges such as supporting collaborative sensemaking process and information literacy education at scale in online communities.

References

[1] Ancestry. 2023. How does Family Tree protect the privacy of living people? Retrieved November 2, 2023 from https://www.familysearch.org/en/help/helpcenter/article/what-is-a-private-space-in-family-tree

[2] Ancestry. 2023. Living People in Family Trees. Retrieved November 2, 2023 from https://support.ancestry.com/s/article/Living-People-in-Family-Trees?language=en_US

[3] Karl Ask and Pär Anders Granhag. 2007. Motivational bias in criminal investigators' judgments of witness reliability 1. *Journal of Applied Social Psychology* 37, 3 (2007), 561–591.

[4] Aadesh Bagmar, Kevin Hogan, Dalia Shalaby, and James Purtilo. 2022. Analyzing the Effectiveness of an Extensible Virtual Moderator. *Proceedings of the ACM on Human-Computer Interaction* 6, GROUP (2022), 1–16.

[5] Aruna D Balakrishnan, Susan R Fussell, Sara Kiesler, and Aniket Kittur. 2010. Pitfalls of information access with visualizations in remote collaborative analysis. In *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. 411–420.

[6] Yasmine Belghith, Sukrit Venkatagiri, and Kurt Luther. 2022. Compete, collaborate, investigate: exploring the social structures of open source intelligence investigations. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–18.

[7] Douglas Belshaw et al. 2012. *What is' digital literacy'? A Pragmatic investigation.* Ph. D. Dissertation. Durham University.

[8] Ronald Bishop. 2008. In the grand scheme of things: An exploration of the meaning of genealogical research. *The journal of popular culture* 41, 3 (2008), 393–412.

[9] Christopher Blomqvist, Kerstin Enflo, Andreas Jakobsson, and Kalle Åström. 2022. Joint handwritten text recognition and word classification for tabular information extraction. In *2022 26th International Conference on Pattern Recognition (ICPR)*. IEEE, 1564–1570.

[10] José Borges. 2019. A contextual family tree visualization design. *Information Visualization* 18, 4 (2019), 439–454.
[11] Wendy Bottero. 2015. Practising family history:'Identity'as a category of social practice. *The British journal of sociology* 66, 3 (2015), 534–556.
[12] Peter Bremer. 2018. Bridging the Gap: Facilitating Genealogical Research. *The Reference Librarian* 59, 1 (2018), 19–22.
[13] Susan L Bryant, Andrea Forte, and Amy Bruckman. 2005. Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia. In *Proceedings of the 2005 ACM International Conference on Supporting Group Work*. 1–10.
[14] Michael Burch, Günter Wallner, Huub van de Wetering, Shahrukh Tufail, Linda Zandt-Sloot, Stasius Gladkis, Minji Hong, and Carlo Lepelaars. 2021. FamSearch: Visual Analysis of Genealogical Data. In *Advances in Visual Computing: 16th International Symposium, ISVC 2021, Virtual Event, October 4-6, 2021, Proceedings, Part II*. Springer, 374–385.
[15] Francesc Calafell and Maarten HD Larmuseau. 2017. The Y chromosome as the most popular marker in genetic genealogy benefits interdisciplinary research. *Human genetics* 136 (2017), 559–573.
[16] Peter Carini. 2016. Information literacy for archives and special collections: defining outcomes. *portal: Libraries and the Academy* 16, 1 (2016), 191–206.
[17] Ruijia Cheng and Benjamin Mako Hill. 2022. Many Destinations, Many Pathways: A Quantitative Analysis of Legitimate Peripheral Participation in Scratch. *Proceedings of the ACM on Human-Computer Interaction* 6, CSCW2 (2022), 1–26.
[18] Victoria Clarke, Virginia Braun, and Nikki Hayfield. 2015. Thematic analysis. *Qualitative psychology: A practical guide to research methods* 3 (2015), 222–248.
[19] Paul Darby and Paul Clough. 2013. Investigating the information-seeking behaviour of genealogists and family historians. *Journal of Information Science* 39, 1 (2013), 73–84.
[20] Jerome De Groot. 2022. Digital Public History: Family History and Genealogy. *Handbook of Digital Public History* (2022), 369.
[21] Hanne De Jaegher and Ezequiel Di Paolo. 2007. Participatory sense-making: An enactive approach to social cognition. *Phenomenology and the cognitive sciences* 6 (2007), 485–507.
[22] Michela Del Vicario, Alessandro Bessi, Fabiana Zollo, Fabio Petroni, Antonio Scala, Guido Caldarelli, H Eugene Stanley, and Walter Quattrociocchi. 2016. The spreading of misinformation online. *Proceedings of the national academy of Sciences* 113, 3 (2016), 554–559.
[23] Wendy Duff and Catherine Johnson. 2003. Where is the list with all the names? Information-seeking behavior of genealogists. *The American Archivist* 66, 1 (2003), 79–95.
[24] Tanya Evans. 2020. Emerging questions in family history studies. *International Public History* 2, 2 (2020), 20190013.
[25] Ixchel Faniel, Eric Kansa, Sarah Whitcher Kansa, Julianna Barrera-Gomez, and Elizabeth Yakel. 2013. The challenges of digging data: a study of context in archaeological data reuse. In *Proceedings of the 13th ACM/IEEE-CS joint conference on Digital libraries*. 295–304.
[26] Board for Certification of Genealogists. 2021. BCG Certification: Who? What? Why? How? Retrieved May 28, 2024 from https://bcgcertification.org/wp-content/uploads/2021/06/BCG-Certification-Brochure-06-2021c.pdf
[27] Board for Certification of Genealogists. 2023. Ethics and Standards. Retrieved September 13, 2023 from https://bcgcertification.org/ethics-standards/
[28] International Commission for the Accreditation of Professional Genealogists. 2023. Why Become an AG® Professional. Retrieved May 28, 2024 from https://www.icapgen.org/become-accredited/
[29] Stephen Paul Forrest III and Tim O Peterson. 2006. It's called andragogy. *Academy of management learning & education* 5, 1 (2006), 113–122.
[30] LS Francis. 2004. The genealogy search process. *PNLA QUARTERLY* 68, 3 (2004), 12–12.
[31] Kate Friday. 2012. *Learning from e-family history: online research behaviour and strategies of family historians and implications for local studies collections.* Ph. D. Dissertation.
[32] Kate Friday. 2014. Learning from e-family history: a model of online family historian research behaviour. *Information Research: An International Electronic Journal* 19, 4 (2014), n4.
[33] Crystal Fulton. 2005. Finding pleasure in information seeking: Leisure and amateur genealogists exploring their Irish ancestry. *Proceedings of the American Society for Information Science and Technology* 42, 1 (2005).
[34] Crystal Fulton. 2009. Quid pro quo: Information sharing in leisure activities. *Library Trends* 57, 4 (2009), 753–768.
[35] Crystal Fulton. 2016. The Genealogist's Information World: Creating Information in the Pursuit of a Hobby. *Journal of Multidisciplinary Research (1947-2900)* 8, 1 (2016).
[36] Elena L Glassman, Aaron Lin, Carrie J Cai, and Robert C Miller. 2016. Learnersourcing personalized hints. In *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing*. 1626–1636.
[37] Claire L Glynn. 2022. Bridging disciplines to form a new one: the emergence of forensic genetic genealogy. *Genes* 13, 8 (2022), 1381.

[38] Kathleen Gregory, Paul Groth, Andrea Scharnhorst, and Sally Wyatt. 2020. Lost or found? discovering data needed for research. *Harvard Data Science Review* (2020).
[39] Jerome de Groot. 2020. Ancestry. com and the evolving nature of historical information companies. *The Public Historian* 42, 1 (2020), 8–28.
[40] Karen Handley, Andrew Sturdy, Robin Fincham, and Timothy Clark. 2006. Within and beyond communities of practice: Making sense of learning through participation, identity and practice. *Journal of management studies* 43, 3 (2006), 641–653.
[41] Derek L Hansen, Patrick J Schone, Douglas Corey, Matthew Reid, and Jake Gehring. 2013. Quality control mechanisms for crowdsourcing: peer review, arbitration, & expertise at familysearch indexing. In *Proceedings of the 2013 conference on Computer supported cooperative work*. 649–660.
[42] J Tuomas Harviainen, Bo-Christer Björk, et al. 2018. Genealogy, GEDCOM, and popularity implications. *Informaatiotutkimus* (2018).
[43] Stephen B Hatton. 2019. History, kinship, identity, and technology: toward answering the question "what is (family) genealogy?". *Genealogy* 3, 1 (2019), 2.
[44] Russell Heimlich. 2012. The '72-Year Rule' Governs Release of Census Records. Retrieved November 1, 2023 from https://www.pewresearch.org/short-reads/2012/04/09/the-72-year-rule-governs-release-of-census-records/
[45] Arnon Hershkovitz and Sharon Hardof-Jaffe. 2017. Genealogy as a lifelong learning endeavor. *Leisure/loisir* 41, 4 (2017), 535–560.
[46] Richards J Heuer. 1999. *Psychology of intelligence analysis*. Center for the Study of Intelligence.
[47] Y Linlin Huang, Kate Starbird, Mania Orand, Stephanie A Stanek, and Heather T Pedersen. 2015. Connected through crisis: Emotional proximity and the spread of misinformation online. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing*. 969–980.
[48] Peng Jiang, Yingrui Yang, Gann Bierner, Fengjie Alex Li, Ruhan Wang, and Azadeh Moghtaderi. 2019. Ranking in genealogy: search results fusion at ancestry. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2754–2764.
[49] S Mo Jones-Jang, Tara Mortensen, and Jingjing Liu. 2021. Does media literacy help identification of fake news? Information literacy helps, but other literacies don't. *American behavioral scientist* 65, 2 (2021), 371–388.
[50] Ruogu Kang, Aimee Kane, and Sara Kiesler. 2014. Teammate inaccuracy blindness: when information sharing tools hinder collaborative analysis. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 797–806.
[51] Enkelejda Kasneci, Kathrin Seßler, Stefan Küchemann, Maria Bannert, Daryna Dementieva, Frank Fischer, Urs Gasser, Georg Groh, Stephan Günnemann, Eyke Hüllermeier, et al. 2023. ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and individual differences* 103 (2023), 102274.
[52] Douglas J Kennard. 2018. Computer-Assisted Crowd Transcription of the US Census with Personalized Assignments for Better Accuracy and Participation. In *Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries*. 41–44.
[53] Heather Kennedy-Eden and Ulrike Gretzel. 2021. My heritage in my pocket: mobile device and app use by genealogy tourists. *Information Technology & Tourism* 23, 3 (2021), 327–350.
[54] M Laeeq Khan and Ika Karlina Idris. 2019. Recognise misinformation and verify before sharing: a reasoned action and information literacy perspective. *Behaviour & Information Technology* 38, 12 (2019), 1194–1212.
[55] Juho Kim et al. 2015. *Learnersourcing: improving learning with collective learner activity*. Ph. D. Dissertation. Massachusetts Institute of Technology.
[56] Nam Wook Kim, Stuart K Card, and Jeffrey Heer. 2010. Tracing genealogical data with timenets. In *Proceedings of the International Conference on Advanced Visual Interfaces*. 241–248.
[57] Soomin Kim, Jinsu Eun, Joseph Seering, and Joonhwan Lee. 2021. Moderator chatbot for deliberative discussion: Effects of discussion structure and discussant facilitation. *Proceedings of the ACM on Human-Computer Interaction* 5, CSCW1 (2021), 1–26.
[58] M.S. Knowles. 1973. *The Adult Learner: A Neglected Species*. Gulf Publishing Company.
[59] Laura Koesten, Kathleen Gregory, Paul Groth, and Elena Simperl. 2021. Talking datasets–understanding data sensemaking behaviours. *International journal of human-computer studies* 146 (2021), 102562.
[60] Laura Koesten, Emilia Kacprzak, Jeni Tennison, and Elena Simperl. 2019. Collaborative practices with structured data: Do tools support what users need?. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–14.
[61] Anne-Marie Kramer. 2011. Mediatizing memory: history, affect and identity in who do you Think you are? *European Journal of Cultural Studies* 14, 4 (2011), 428–445.
[62] Lisa Krizan. 1999. *Intelligence essentials for everyone*. Number 6. Joint Military Intelligence College.

[63] Tiffany H Kung, Morgan Cheatham, Arielle Medenilla, Czarina Sillos, Lorie De Leon, Camille Elepaño, Maria Madriaga, Rimel Aggabao, Giezel Diaz-Candido, James Maningo, et al. 2023. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLoS digital health* 2, 2 (2023), e0000198.
[64] Ronald D Lambert. 2002. Reclaiming the ancestral past: Narrative, rhetoric and the 'convict stain'. *Journal of Sociology* 38, 2 (2002), 111–127.
[65] Jean Lave and Etienne Wenger. 1991. *Situated learning: Legitimate peripheral participation.* Cambridge university press.
[66] Tianyi Li, Chandler J Manns, Chris North, and Kurt Luther. 2019. Dropping the baton? Understanding errors and bottlenecks in a crowdsourced sensemaking pipeline. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (2019), 1–26.
[67] Yuhua Liu, Sicheng Dai, Changbo Wang, Zhiguang Zhou, and Huamin Qu. 2017. GenealogyVis: A system for visual analysis of multidimensional genealogical data. *IEEE Transactions on Human-Machine Systems* 47, 6 (2017), 873–885.
[68] Thomas P Mackey and Trudi E Jacobson. 2011. Reframing information literacy as a metaliteracy. *College & research libraries* 72, 1 (2011), 62–78.
[69] Eric Malmi, Aristides Gionis, and Arno Solin. 2018. Computationally inferred genealogical networks uncover long-term trends in assortative mating. In *Proceedings of the 2018 World Wide Web Conference*. 883–892.
[70] Eric Malmi, Marko Rasa, and Aristides Gionis. 2017. AncestryAI: A tool for exploring computationally inferred family trees. In *Proceedings of the 26th International Conference on World Wide Web Companion*. 257–261.
[71] Yazdan Mansourian. 2021. Joyful information activities in serious leisure: looking for pleasure, passion and purpose. *Aslib Journal of Information Management* 73, 5 (2021), 601–617.
[72] Tamara Marksteiner, Karl Ask, Marc-André Reinhard, and Pär Anders Granhag. 2011. Asymmetrical scepticism towards criminal evidence: The role of goal-and belief-consistency. *Applied Cognitive Psychology* 25, 4 (2011), 541–547.
[73] Elizabeth S Mills. 1999. Working with historical evidence: Genealogical principles and standards. *National Genealogical Society Quarterly* 87, 3 (1999), 165–84.
[74] Vikram Mohanty, David Thames, Sneha Mehta, and Kurt Luther. 2020. Photo sleuth: Identifying historical portraits with face recognition and crowdsourced human expertise. *ACM Transactions on Interactive Intelligent Systems (TiiS)* 10, 4 (2020), 1–36.
[75] Susan M Moore and Doreen A Rosenthal. 2021. What Motivates Family Historians? A Pilot Scale to Measure Psychosocial Drivers of Research into Personal Ancestry. *Genealogy* 5, 3 (2021), 83.
[76] Sunny Jane Morton. 2019. The World's Largest Shared Family Tree. Retrieved September 13, 2023 from https://www.familysearch.org/en/blog/online-family-tree
[77] Christian Murphy, Gail Kaiser, Kristin Loveland, and Sahar Hasan. 2009. Retina: helping students and instructors based on observed programming activities. In *Proceedings of the 40th ACM technical symposium on Computer Science Education*. 178–182.
[78] Catherine Nash. 2004. Genetic kinship. *Cultural studies* 18, 1 (2004), 1–33.
[79] National Academies of Sciences, Engineering, and Medicine. 2018. How people learn II: Learners, contexts, and cultures.
[80] Jessica Neaves. 2021. 7 Simple Steps To Creating A Genealogy Timeline (And Why You Need One). Retrieved January 08, 2024 from https://www.heritagediscovered.com/blog/7-simple-steps-creating-genealogy-timeline-why-need-one
[81] Johnny Nhan, Laura Huey, and Ryan Broll. 2017. Digilantism: An analysis of crowdsourcing and the Boston marathon bombings. *British Journal of Criminology* 57, 2 (2017), 341–361.
[82] Thibauld Nion, Farès Menasri, Jérôme Louradour, Cédric Sibade, Thomas Retornaz, Pierre-Yves Métaireau, and Christopher Kermorvant. 2013. Handwritten information extraction from historical census documents. In *2013 12th International Conference on Document Analysis and Recognition*. IEEE, 822–826.
[83] Association of College and Research Libraries. 2000. Information literacy competency standards for higher education. (2000).
[84] Association of College and Research Libraries. 2016. Framework for information literacy for higher education. (2016).
[85] Thomas C Ormerod, Emma Barrett, and Paul J Taylor. 2008. Investigative sense-making in criminal contexts. *Naturalistic decision making and macrocognition* (2008), 81–102.
[86] Chris Piech, Mehran Sahami, Daphne Koller, Steve Cooper, and Paulo Blikstein. 2012. Modeling how students learn to program. In *Proceedings of the 43rd ACM technical symposium on Computer Science Education*. 153–160.
[87] Peter Pirolli and Stuart Card. 2005. The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis. In *Proceedings of international conference on intelligence analysis*, Vol. 5. McLean, VA, USA, 2–4.
[88] Solène Prince. 2022. Affect and performance in ancestral tourism: stories of everyday life, personal heritage, and the family. *Journal of Heritage Tourism* 17, 1 (2022), 20–36.

[89] Antonio Regalado. 2019. More than 26 million people have taken an at-home ancestry test. Retrieved January 16, 2024 from https://www.technologyreview.com/2019/02/11/103446/more-than-26-million-people-have-taken-an-at-home-ancestry-test/
[90] Joseph P Robinson, Zaid Khan, Yu Yin, Ming Shao, and Yun Fu. 2021. Families in wild multimedia: A multimodal database for recognizing kinship. *IEEE Transactions on Multimedia* 24 (2021), 3582–3594.
[91] Christine Rose. 2014. *Genealogical Proof Standard: Building a Solid Case* (4th ed.). CR Publications.
[92] Henriette Roued, Helene Castenbrandt, and Bárbara Ana Revuelta-Eugercios. 2022. Search, save and share: family historians' engagement practices with digital platforms. *Archival Science* (2022), 1–20.
[93] Barış Serim and Giulio Jacucci. 2019. Explicating" Implicit Interaction" An examination of the concept and challenges for research. In *Proceedings of the 2019 chi conference on human factors in computing systems*. 1–16.
[94] Farhana Shahid, Srujana Kamath, Annie Sidotam, Vivian Jiang, Alexa Batino, and Aditya Vashistha. 2022. " It Matches My Worldview": Examining Perceptions and Attitudes Around Fake Videos. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–15.
[95] Aviv J Sharon and Ayelet Baram-Tsabari. 2020. Can science literacy help individuals identify misinformation in everyday life? *Science Education* 104, 5 (2020), 873–894.
[96] Emma Shaw. 2020. "Who we are, and why we do it": a demographic overview and the cited motivations of Australia's family historians. *Journal of Family History* 45, 1 (2020), 109–124.
[97] National Genealogical Society. 2018. Becoming a Professional Genealogist. Retrieved May 28, 2024 from https://www.ngsgenealogy.org/free-resources/becoming-a-professional-genealogist/
[98] National Genealogical Society. 2018. NGS Guidelines. Retrieved September 14, 2023 from https://www.ngsgenealogy.org/free-resources/guidelines/
[99] Philipp Spitzer, Niklas Kühl, Daniel Heinz, and Gerhard Satzger. 2023. ML-Based Teaching Systems: A Conceptual Framework. *arXiv preprint arXiv:2305.07681* (2023).
[100] Matthew Stallard and Jerome De Groot. 2020. "Things are coming out that are questionable, we never knew about": DNA and the new family history. *Journal of family history* 45, 3 (2020), 274–294.
[101] Kate Starbird, Jim Maddock, Mania Orand, Peg Achterman, and Robert M Mason. 2014. Rumors, false flags, and digital vigilantes: Misinformation on twitter after the 2013 boston marathon bombing. *IConference 2014 proceedings* (2014).
[102] United Nations Educational, Scientific and Cultural Organization. 2023. Information Literacy. Retrieved November 13, 2023 from https://www.unesco.org/en/ifap/information-literacy
[103] Ricardo Urrestarazu-Capellán, José Correa-Fernández, and Francisco Sánchez-Cubo. 2022. The sources of information of the genealogical tourist: the influence of social networks and genealogical associations. *Heliyon* (2022), e11551.
[104] Kylie Veale. 2004. Discussing our family trees: A longitudinal analysis of online, community-based communication in genealogical newsgroups. *Interface on the Internet* 4, 4 (2004).
[105] Kylie Veale. 2006. Genealogical education: Finding Internet-based educational content for hobbyist genealogists. In *The international handbook of virtual learning environments*. Springer, 939–959.
[106] Fernanda B Viégas, Martin Wattenberg, and Matthew M McKeon. 2007. The hidden order of Wikipedia. In *Online Communities and Social Computing: Second International Conference, OCSC 2007, Held as Part of HCI International 2007, Beijing, China, July 22-27, 2007. Proceedings 2*. Springer, 445–454.
[107] Nai-Ching Wang, David Hicks, and Kurt Luther. 2018. Exploring trade-offs between learning and productivity in crowdsourced history. *Proceedings of the ACM on Human-Computer Interaction* 2, CSCW (2018), 1–24.
[108] Ruan Wang, Jun Deng, Xinhui Guan, and Yuming He. 2023. A framework of genealogy knowledge reasoning and visualization based on a knowledge graph. *Library Hi Tech* (2023).
[109] Etienne Wenger. 1999. *Communities of practice: Learning, meaning, and identity*. Cambridge university press.
[110] Etienne Wenger. 2011. Communities of practice: A brief introduction. (2011).
[111] FamilySearch Wiki. 2023. FamilySearch Research Wiki. Retrieved January 05, 2024 from https://www.familysearch.org/en/wiki/Main_Page
[112] FamilySearch Wiki. 2023. Research Logs. Retrieved Janurary 02, 2024 from https://www.familysearch.org/en/wiki/index.php?title=Research_Logs
[113] Wikipedia. 2023. GEDCOM. Retrieved Janurary 03, 2024 from https://en.wikipedia.org/wiki/GEDCOM
[114] Jaci Wilkinson and Natalie Bond. 2021. Digital Literacy Skills for Family History Research. *Evidence Based Library and Information Practice* 16, 2 (2021), 89–110.
[115] Heather Willever-Farr, Lisl Zach, and Andrea Forte. 2012. Tell me about my family: A study of cooperative research on Ancestry. com. In *Proceedings of the 2012 iConference*. 303–310.
[116] Heather L Willever-Farr and Andrea Forte. 2014. Family matters: Control and conflict in online family history production. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*. 475–486.
[117] Tom D Wilson. 1999. Models in information behaviour research. *Journal of documentation* 55, 3 (1999), 249–270.

[118] Thomas D Wilson. 2000. Human information behavior. *Informing science* 3 (2000), 49.
[119] Elizabeth Yakel. 2004. Seeking Information, Seeking Connections, Seeking Meaning: Genealogists and Family Historians. *Information Research: an international electronic journal* 10, 1 (2004), n1.
[120] Elizabeth Yakel and Deborah Torres. 2007. Genealogists as a" Community of Records". *The American Archivist* 70, 1 (2007), 93–113.
[121] Ashley Ge Zhang, Yan Chen, and Steve Oney. 2023. VizProg: Identifying Misunderstandings By Visualizing Students' Coding Progress. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–16.
[122] Pengyi Zhang and Dagobert Soergel. 2014. Towards a comprehensive model of the cognitive process and mechanisms of individual sensemaking. *Journal of the Association for Information Science and Technology* 65, 9 (2014), 1733–1756.
[123] Ann Zimmerman. 2007. Not by metadata alone: The use of diverse forms of knowledge to locate data for reuse. *International Journal on Digital Libraries* 7 (2007), 5–16.

A Interview Guide

Introduction:

(1) How do you become interested in genealogy?
  (a) What is your occupation? Is it related to genealogy?
(2) Can you tell me more about your research?
  (a) What is the primary focus (own family, others', research as a professional, etc.)?
  (b) How long have you been doing it?
  (c) How often are you involved in genealogy related activities?
  (d) (Non-credential) How would you describe your expertise on genealogical research (novices, intermediate, expert, etc.)?
(3) What online genealogy communities do you participate in (Facebook, Reddit, Email lists, etc.)? How do you use them (seeking help, offering help, seek information, etc.)?
(4) (If tools include DNA) How do you use DNA test results for your genealogy research? (If not) How do you feel about the DNA test? would you consider using them for your research?

Ongoing research:

(5) How do you decide which question to start with or focus on in your research (what's easiest, what's most interesting, what the client asked for, etc.)?
(6) Can you tell me about a time when you went down a research path or followed a hunch that you weren't sure was correct or not?
(7) Have you ever explored several different answers to the same question simultaneously?
(8) How would you determine when a genealogy question is answered?
(9) How much time do you usually expect to solve a genealogy problem or decide to move to another topic?
(10) How do you keep track of the progress before it's solved?
(11) Can you tell us about one time when you disagreed with other people's conclusion? How did you resolve it?

Research barriers:

(12) Can you tell me about a time when you ran into a brick wall?
(13) What would you do to avoid or overcome brick walls?
(14) (Non-credential) How do you feel about credentials in genealogy research? Would you consider hiring a professional genealogist to research for you?
(15) (Non-credential) Where do you seek help in your research? How do you tell someone is good at genealogy?

Information:

(16) How reliable do you think the records provided by websites like Ancestry.com and FamilySearch.org? How about the indexes?
(17) How reliable are the record hints provided in these websites? How do you use them in your research?
(18) How reliable do you think user-created family trees are? How do you use them in your research?
(19) (If negative) What do you think are the causes of this issue?
(20) (If negative) How do you identify and avoid unreliable information?

Sharing:

(21) How do you share your research outcome with others?
  (a) What steps would you take? In what formats?
  (b) What would you do to ensure the accuracy of your work?

Education:

(22) How did you learn how to do genealogy research?
(23) Are there any tools that you found could help newcomers?
  (a) (Expert) How can you recognize a newcomer on the Internet?
(24) Can you imagine how we may help people become better genealogy researchers?

Changes:

(25) (For experienced genealogists) What do you think about genealogy have changed in the past decade?

Reflection:

(26) Any additional thoughts before we wrap up for today? Any other aspect of your research we should have asked but didn't?
(27) Anything you want to ask us?

Received January 2024; revised July 2024; accepted October 2024