Trust in Qualitative Data Repositories

Rebecca D. Frank, Zui Chen, Erica Crawford, Kara Suzuka, Elizabeth Yakel

ABSTRACT

This paper applies Pirson and Malhotra's (2011) framework for organizational trust to data reuse in an educational context. Their framework delineates stakeholders along two axes: depth of the interactions (e.g., shallow or deep) and the locus of the relationship (e.g., internal or external). We analyzed 139 survey responses and 44 in-depth interviews with users of repositories holding video records of practice. We found that factors such as data quality, co-production of data between the repository and data producer, responsiveness of staff, and transparency of curation processes influenced trust, and that responsiveness and transparency were of particular importance for trust development for users with deeper interactions and a more internal loci of relationships to repositories.

Keywords

Data Reuse, Digital Curation, Digital Repositories, Qualitative Data, Trust

Rebecca D. Frank
School of Information, University of Michigan, 3429 North Quad, 105 S. State St., Ann Arbor, MI 48109-1285;
frankrd@umich.edu

Zui Chen
School of Information, University of Michigan, 3429 North Quad, 105 S. State St., Ann Arbor, MI 48109-1285;
zuic@umich.edu

Erica Crawford
School of Education, University of Michigan, 10 E University Ave., Ann Arbor, MI 48109;
zawackie@umich.edu

Kara Suzuka
College of Education, University of Hawai'i at Mānoa, 1776 University Ave., Honolulu, Hawai'i 96822;
klms@hawaii.edu

Elizabeth Yakel
School of Information, University of Michigan, 3429 North Quad, 105 S. State St., Ann Arbor, MI 48109-1285;
yakel@umich.edu

INTRODUCTION

Data reuse is a practice that relies on trust among data reusers, data producers, data repositories, and the data itself. Trust in this context is a multidimensional concept. Reusers' decisions about data selection and reuse are influenced by the depth of their interactions and positioning of their relationships to the data producer, repository, and data (Pirson & Malhotra, 2011). The depth of reuse (extent and intensity) and locus (internal or external) of relationships come into play when considering how trust influences the reuse of video records of practice in education. Video records of practice consist of raw or produced footage of actual teaching and/or learning activities in a classroom. Some document specific events and last a few minutes; others last for hours and document an entire lesson or set of lessons. Video records of practice are a key type of data used in educational research and are also used in pre-service and continuing teacher education (Bass et al., 2002). Because video records of practice are costly, time consuming, and can be challenging to produce, researchers and educators increasingly rely on repositories to curate, preserve, and provide access to these data. The organizations that provide video records of practice for reuse range from traditional data repositories to diverse education organizations serving the profession, institutions, and individuals.

Given the range of organizations involved in curating video records of practice, we were interested in exploring issues of trust in the repositories and the data. Previous studies have focused on more traditional data and repositories and trust (Donaldson & Conway, 2015; Yakel, Faniel, Kriesberg, & Yoon, 2013) or research data that is more numerical or textual (Yoon, 2014). In the management information systems literature, trust in institutions is linked to data quality, but this research centers on for-profit organizations (Lee, Strong, Kahn, & Wang, 2002).

In this paper, we apply Pirson and Malhotra's (2011) framework for organizational trust to data reuse in an education context. The framework delineates stakeholders along two axes: depth of the interaction with the data and the institution (e.g., shallow or deep) and the locus of the relationship (e.g., internal or external) (2011, p. 1091). In their study, Pirson and Malhotra also identified transparency as a dimension of trust in organizations, in cases where stakeholders have deep, internal relationships to organizations (2011, p. 1097). This framework provides an interesting basis for examining trust in the data reuse process. We defined depth of interaction as the extent or intensity of data reuse and it fell along a continuum from those who engaged in less extensive reuse, such as reusing a single, targeted video to illustrate an approach to teaching fractions, to reusers deeply engaged in highly intense research reuse who utilize multiple videos of particular teachers over time. We analyzed the relationship with the repository dimension along a continuum ranging from reusers who had worked for the repository to those who knew little about the repository and may only have interacted with the website to access digital records of practice.

We extend previous studies by asking whether interactions between data reusers and repository staff influence attitudes about, and trust in, repositories. The notion that repositories can better meet the needs of their users through sustained interactions is consistent with Kelton's argument that "Trust does not flow from a trustor to a trustee; rather, it emerges from their interaction over time" (Kelton, Fleischmann, & Wallace, 2008, p. 364).

Our study is motivated by the following research questions:

- What is the number and range of repositories with which data reusers interact?

- How can we assess the depth of interactions that data reusers have with repositories and repository staff?

- How do data reusers' relationships to repositories along an internal / external continuum influence their behaviors and attitudes toward repositories?

- When and how do depth of interaction and/or relationship with the repository influence data reusers' trust in digital repositories and the data that they provide?

We report on the results of a mixed methods study involving 139 survey responses from users of three repositories and 44 in-depth interviews with reusers of video records of practice. We focused on data reusers' interactions with repository staff, relationships with repositories, and how those interactions influenced their attitudes about repositories, including trust in the repository and trust in the data that it curates and makes accessible.

LITERATURE

Data Reuse

Researchers engaged in data reuse or secondary use of research data often have different goals than those of the original data producer (Corti, 2000; Karasti & Baker, 2008; Yoon, 2016; Zimmerman, 2008). In the field of education, data reusers have three main goals for their reuse: research, teaching, and professional development (e.g. Derry et al., 2010; Gaudin & Chaliès, 2015; Gonzalez, Deal, & Skultety, 2016). Therefore, data originally created for research can be used not only to answer different research questions, but also for the entirely different end goal of professional education.

Data reuse depends in part on data sharing, and in order for data to be shareable it must be both searchable and usable (Poole, 2015; Schofield et al., 2009). Data sharing and reuse are challenging even when best practices for data management are followed (Akmon, Zimmerman, Daniels, & Hedstrom, 2011; Poole, 2015). Qualitative data in particular presents unique challenges, relating to the need for sufficient contextual information to facilitate the work of data reusers (Corti, 2000; Mauthner & Parry, 2009).

Although many of the unique challenges of providing this contextual information are put upon the data producers, data sharing and reuse are facilitated by the people and technologies that form the data curation infrastructure (e.g. Bishop, 2012; Donnelly & North, 2011; Poole, 2015; Pryor & Donnelly, 2009). Digital repositories, as part of that infrastructure, support data producers and reusers throughout the data lifecycle: from helping data producers develop data management plans at the outset of their research, to supporting data management during a project, and encouraging preservation and facilitating data reuse.

Data Repositories, Data Curation, Repository Staff

In addition to preserving, curating, and providing access to data (Corti, 2011; Giarlo, 2013), one of the primary aims of data repositories is "to facilitate communication, collaboration and interactions" among data producers and reusers (Chowdhury, 2010). The responsibility to meet the needs of repository reusers beyond mere access to data (Russell & Day, 2010; Yoon, 2014) may be in conflict with some of the ways that repositories are structured.

For example, Margaryan and Littlejohn (2008) argued that repositories are designed for the capabilities of technology rather than the needs of users. In a study examining the sharing and reuse of digital learning resources, they found that repository managers were focused on long-term preservation, and repository-centric data management activities to support those long-term goals, while users focused on short-term outcomes and a wider range of contextual factors to meet their immediate needs. They also found that interoperability issues with repository tools and users' technology was a barrier to repository use (Margaryan & Littlejohn, 2008). Similarly, in a previous study we found that specialized repository tools for working with data were a barrier to reuse of digital records of practice in education (Frank, Suzuka, & Yakel, 2016). Russell and Day (2010) have also noted the disconnect between

repository goals and user needs. Their findings suggest that increased interaction between repository staff and users can improve user experience and engagement with repositories.

Trust in Repositories

We use a broad definition of repository as "any type of organization that holds documents, including business, institutional, and government archives, manuscript collections, libraries, museums, and historical societies, and in any form, including manuscripts, photographs, moving image and sound materials, and their electronic equivalents . . . where things can be stored and maintained."[4] We focus on digital repositories where, "contemporary understanding has broadened from an initial focus on software systems to a wider and overall commitment to the stewardship of digital materials; this requires not just software and hardware, but also policies, processes, services, and people, as well as content and metadata."[5]

We also use the definition of trust established by Yoon, "Trust is the data reusers' belief that the data will result in positive outcomes, leading to the reuse of such data in their research. Data reusers' trust judgments can be understood as psychological processes, and whether they accept and use certain data can be seen as an indication of trusting behavior" (2016, p. 2). Researchers can engage in different levels of trust throughout the process of data reuse in response to a variety of factors, such as data quality, data completeness, curation, and data management practices (Yoon, 2016).

The varying nature of the trust relationship has also been explored by Bak (2016), who described trust in the context of digital repositories as both social and contingent, embedded within relationships. This view of trust is pertinent to our study, which examines how interaction with data and the repository staff influences attitudes about, and trust in, digital repositories.

The concept of trustworthiness for digital repositories is based on a perspective that relies on research from both archival science and web credibility (Donaldson & Conway, 2015). Examining trust from these two perspectives provides an understanding from both the perspective of repository staff as well as users. In addition to trust in digital repositories, trust between data producers, repository staff, and data reusers is foundational for data sharing and reuse (e.g. Akmon et al., 2011; Faniel & Jacobsen, 2010; Poole, 2015; Yakel et al., 2013). Yakel et al. (2013) found that reuser trust in digital repositories is linked to repository processes. In this paper, we argue that interaction and communication between repository staff and data reusers can help reusers to understand repository processes for curating and managing qualitative data in education.

Digital Records of Practice in Education

Qualitative data in the field of education frequently consists of records of practice. Records of practice are "detailed documentation of teaching and learning" (Bass et al., 2002, p. 79). They generally depict the work of teaching and learning that takes place in a classroom environment, and can include video, audio, still images, lesson plans, student work, and transcripts. Digital records of practice, such as classroom video, present curation and reuse challenges for repositories as well as data reusers.

We focus here on video records of practice, which have a long history of use in the field of education for research, teaching, and professional development (e.g. Burleigh & Peterson, 1967; Shelton, Archambault, Warren, Ahmad, & Nevárez, 2016; Stigler, Gallimore, & Hiebert, 2000). One important way that video records of practice are used in teacher education and professional development

---

[4] http://archivists.org/glossary/terms/r/repository

[5] http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/digital-repositories

is to help pre-service and in-service teachers develop professional vision and noticing skills (e.g. Penn-Edwards, 2012; Rook & McDonald, 2012; van Es & Sherin, 2002). For research, video records of practice enable multiple researchers to view and re-view classroom events (Hadfield & Haw, 2012).

Despite these benefits and affordances, Pea has argued that "there is a serious and persistent gap between such promise and the usefulness of video records" (Pea, 2006, p. 1325). Some of the issues that persist include problems with capture, storage, and sharing of video files (e.g. Zhang, Lundeberg, Koehler, & Eberhardt, 2011).

RESEARCH METHODS

This paper is part of a larger research project, Qualitative Data Reuse: Records of Practice in Educational Research and Teacher Development (QDR), which examines the reuse of qualitative video data in education through a mixed methods approach utilizing surveys, interviews, web analytics, and observations. The results in this paper draw upon the survey and interview results. The Institutional Review Board at the primary author's university reviewed the research design and deemed it exempt.

Interviews

We conducted 44 interviews with educational researchers and both pre-service and in-service teacher educators who reuse digital video. We used a convenience sampling technique to recruit initial interviewees. This included recruiting data reusers who we identified through the research literature, at professional conferences (i.e. those presenting on their work with video records of practice), and by employing a snowball sampling technique that asked interviewees to recommend others. Of the 44 interviewees, 31 identified themselves as university faculty, five as graduate students, two as postdoctoral researchers, and six as other types of education professionals.

We focused this analysis on the 21 interviewees who said that they had interacted with repository staff. Of those 21 interviewees, 11 identified as faculty, five as graduate students, one as a postdoctoral researcher, and 4 as other types of education professionals. These interviewees also identified their primary area of data reuse, with 11 reporting their primary reuse for research, eight for pre-service teacher training, one for teacher professional development, and one for personal study. Although 17 of the 21 interviewees identified as data producers as well as reusers, only five said that they have deposited their own data in a repository.

We asked the interviewees about their data reuse practices, including questions about interactions with repository staff, issues around trust, and attitudes about digital repositories. Each semi-structured interview lasted approximately one hour, and was audio-recorded and then transcribed for analysis.

We analyzed the transcripts using NVivo, a qualitative data analysis software. Our codeset was based on themes from the literature as well as themes that arose during the coding and analysis (Miles & Huberman, 1994; Saldaña, 2015). After a series of paired coding on the same transcripts, the coders reached a reliability rating of 0.712 (for interviews focused primarily on research) and 0.732 (for those focused on teaching) using Scott's Pi, a statistic that measures inter-rater reliability for coding textual data (Holsti, 1969).

After the first round of coding, we used axial coding to examine the data further, focusing on the following codes: interaction with repository staff, attitudes about repositories, and trust in repositories.

Survey

We developed the survey based on preliminary interviews, a review of the literature, and cognitive walkthroughs with preliminary versions of the survey itself. The survey is divided into

sections focusing on reuse goals (i.e. research, teaching, and personal study), as well as questions about general reuse, and participant demographics.

We received 139 responses from the users of three repositories. Our respondents included students, faculty, and researchers who reported using digital records of practice for research, teaching, and personal study.

We analyzed survey responses using SPSS, a statistical analysis software package. For this article, we focused specifically on questions about interaction with repository staff and participant attitudes about repositories.

FINDINGS

Our findings are organized into three sections. First, we examine data reuser interaction with repositories, including the types of interaction with repository staff. Second, we discuss data reuser attitudes about repositories. And finally, we explore data reusers' trust in digital repositories.

Interaction with Repositories

Survey respondents were asked to report the repositories they used to find digital records of practice (Table 1). Data reusers were permitted to identify multiple repositories.

Repository	Type	Number of Data Reusers
Accomplished Teaching, Learning and Schools (ATLAS)	Professional Organization	27
The Inter-university Consortium for Political and Social Research (ICPSR)	University-based Video Archives	2
Teaching and Learning Exploratory (TLE)	University-based Video Archives	33
Teaching Channel	Non-profit Organization	44
Trends in International Mathematics and Science Study Video (TIMSS)	University-based Video Archives	22
Virtual Learning Community (VLC)	University-based Video Archive	20
Not Sure		3
Other		26

Table 1: Repository Usage (n=84)

Respondents reported interactions with a range of organizations serving as repositories for video records of practice. On average, they used 2.1 repositories. Twenty-two respondents (26.2%) used a single repository, 40 (47.6%) used two repositories, 14 (16.7%) used three repositories, and seven (8.3%) used four repositories. One respondent reported using five, the maximum number of repositories named in the survey. Due to the survey design, all respondents to this question reported using at least one repository.

Of the repositories listed in Table 1, one is managed by professional organization, one by a not-for-profit organization, and four are university-based video collections. ATLAS[6] is managed by the National Board for Professional Teaching Standards, and is meant to illustrate and support the work of Board-certified teachers. Teaching Channel[7] was established as a grant-funded, non-profit providing an interactive platform where teachers can use video and other records of practice for learning and

---

[6] http://www.nbpts.org/atlas

[7] https://www.teachingchannel.org

development. ICPSR[8], TLE[9], TIMSS[10], and VLC[11] are all university-based video archives that provide access to videos and other records of practice for research, teaching, and professional development.

We found that three respondents were not sure about the source of the videos that they have used, and that 26 respondents reported using at least one repository that was not listed in our survey.

We also asked participants how frequently they used digital records of practice. Of the 130 respondents who answered this question, 31 (23.8%) said that they used digital records of practice frequently and that they were a major basis for their research or teaching; 85 respondents (65.4%) used digital records of practice occasionally, for a particular class topic or research purpose; and 14 respondents (10.8%) used digital records of practice rarely.

Interaction with Repository Staff

Findings from both the survey and the interviews indicated that few data reusers interacted with staff from any repository. Those who did interact did so for a variety of reasons, including incidental interaction at professional conferences or workshops, email correspondence requesting help, interaction via colleagues or team members, and previous experience working with or for a repository. These types of interactions and relationships with the repository show the range of depth of interaction as well as the closeness of affiliation (internal/external) suggested by Pirson and Malhotra (2011).

We asked our survey respondents if they had interacted with repository staff members when reusing digital records of practice from a repository. Of the 122 respondents who answered this question, a majority (105 respondents, or 86.1%) had not; only 17 respondents (13.9%) reported seeking assistance from repository staff when using digital records of practice. The lack of staff interaction is not surprising. Anecdotally, during the cognitive interviews for the survey, one participant expressed surprise that one could even reach out to a staff person through a digital repository.

Though the survey questions did not explicitly quantify participants' reasons for their interactions with repository staff, one question probed for the types of assistance they requested and yielded 12 responses. Nine needed technical assistance or help troubleshooting software applications provided by the repository, and four reported requesting help with navigation and/or questions about the resources themselves. One respondent asked about further development of repository services for increased usability.

The respondents who reported interacting with repository staff were further asked to assess the importance of the assistance that they received. Of the 17 respondents who were given this question, 15 reusers (88.2%) indicated that staff were either "very helpful" or "essential" to their activities. Only two respondents described repository staff as being "somewhat helpful," and none responded that they were unhelpful.

The interviewees provided more in-depth discussions of the nature of repository staff assistance. Two interviewees described readily available support from repositories in using their data; both also noted that the help was invaluable in facilitating their reuse of records of practice. Interviewee 029 described attending a repository workshop about using a particular dataset and emphasized that this

---

[8] https://www.icpsr.umich.edu

[9] https://tle.soe.umich.edu

[10] http://www.timssvideo.com

[11] https://vlc.uchicago.edu

intensive, in-person interaction continued to be helpful for years: "I went to the workshop at [repository], maybe almost two years ago at this point, which is really helpful in orienting to that data." Interviewee 029 is interesting because her discussion demonstrates that although her reuse might be shallow in terms of the quantity of videos used, it was critical to her research, so intensity may be another way to measure the depth of the interaction.

Interviewee 037 had been using videos from the same repository for several years, and the help she received eliminated challenges that interfered with her ability to reuse video records of practice, "I've been with them for a couple years now, and any time there's been a challenge, I call them up and say, 'Can you help me fix this?'" She continued, "They're pretty quick at responding, too. I have three points of contact in the office . . . I had a technical problem one time, and I could get on the site but the video wasn't working, and [name of repository] was down. They got working on it, and like two minutes before my class began, they got it back up and running. I call them in an emergency. It's usually last minute."

Interviewees 002, 006, 008, 014, 040, and 041 all explained that they knew the staff in the repositories because they had previously worked with or for those repositories, "I started doing some consulting work for the [organization] on their video archives" (002). These reusers represent a group with both intensive and deep use of the resources and stronger internal ties to the repositories they used. Repository processes were also more transparent to them than to reusers with less experience working in and directly with repositories.

Interaction with repository staff was also facilitated by colleagues. Interviewee 022 said that her research team was working directly with the staff of a repository, and that the partnership came about through one of her team members, "And then now that [repository] is up and running, we're actually partnering with them, for a different research study, not the one I just described."

Both survey respondents and interviewees described interactions with repository staff across diverse activities and contexts. These interactions demonstrate both data use and repository relationships of varying depth and breadth. In many cases repository staff provided support or assistance to facilitate the reuse of digital records of practice, and, when this occurred, the interactions were generally described in positive terms by the reusers.

Attitudes about Repositories

We next examined the attitudes about digital repositories, in light of participant interaction with repositories and repository staff. We found that a majority of survey respondents would use data from a repository again. Interviewees who interacted with repository staff had mixed feelings about using a repository again, with positive, negative, and ambivalent or uncertain attitudes.

Our survey asked data reusers about their intentions to again reuse digital records of practice: "I would use digital records of practice from third party repositories or organizations for future projects." This was asked of all respondents, regardless of whether they had reported interaction with repository staff or not. Of the 126 respondents who answered this question, 115 (91.3%) answered positively that they would use digital records of practice from a repository again in the future. Of the 17 respondents who reported interacting with repository staff, all (100%) affirmed that they expected to use digital records of practice from repositories in the future.

Survey respondents indicating they would use repository data again in the future were given the option to include their reasoning. Of the 17 respondents who had both requested assistance from repository staff and expect to use digital records of practice in the future, nine (52.9%) provided this

additional input. These responses generally emphasized the importance of access to a large variety of videos as well as the relevance of the videos to their work.

Interviewees 005, 009, and 022 were less closely associated to the repositories they used. Each expressed positive views about how digital repositories curated the collections. For example, 005 talked about the usefulness of repositories and the curation work by repository staff, "I think, that as a teacher-educator the ones that are pretty curated are often useful because you often are trying to look for really specific types of things. If I wasn't involved in [repository], I can't really imagine wanting to sit there and watch all of that video to try to find whether or not there were things I could use in my math methods class."

Similarly, 009 noted the added value of highly curated repository collections for teacher-educators using video records of practice for their own professional learning, "if you are just on your own, like me, where you're not in a PD [professional development], you don't have anybody working with you, you don't have anybody facilitating, you don't have anything, that when you find something, what you watch something on [repository], they have all those questions, the scaffolding on the side . . . And also the way they have the documents attached." She also explained that she appreciates the predictability that comes with searching for records of practice in a highly curated repository collection, "because they're all such high quality that happens in a very predictable way, there's always something else good to watch."

Interviewee 022 emphasized the value and importance of repositories making video records of practice available and findable in her field, "There could be amazing video out there, but if I can't find it, then it doesn't help me, you know? So if somebody is making it available and findable, I think that is a huge value add."

In all three cases, interviewees found that the work of repository staff to make collections usable and accessible facilitated added value to their own work with video records of practice. Interviewee 017 framed this value in broad terms, as a service to the field of education, "I think it's a tremendous and important service to the field." When asked if those repositories added value to the records of practice that they curate, she was similarly positive and appreciated the transparency of their actions. "I haven't had tons of experience with it . . . but I do know like the more thoughtfully done it is done, the more useful it is to the field. And so I've had some interaction with folks at [institution]…around the repository that they're trying to develop and have just really appreciated some of the detail of like how they're trying to figure it out and how to make it accessible."

Some interviewees tempered their positive views of repositories with cautionary statements about either the completeness of collections or curation work that appeared to be pushing an agenda. Interviewee 018 explained that in addition to making video records of practice accessible, repositories also "provide some kind of context" which adds value to the records of practice. However, she cautioned that this kind of work is only helpful if repositories are "doing something in a very neutral way . . . but if they have some kind of hidden agenda and they are kind of making things look a certain way" then she thought of the repository as less trustworthy and the work of repository staff less valuable.

Several interviewees were more critical in their views about digital repositories. Interviewee 002 had experience working as a consultant for a repository, which made her skeptical about the availability and/or quality of support that she could expect from repositories generally, "I just was doubtful they would respond if I wrote… It's writing the national government. Like, who's really gonna respond? …But I know that people reach out when they're trying to use the [repository]. When we were trying to tag in the [repository], there were multiple emails that got sent out. And the most frustrating thing is there was no response and we were working for them." After the conclusion of her work for the

repository she reached back out to the staff with questions about using the data, but "there was never ever any follow up and I was just done." In this case, a deep, internal relationship with a repository increased her desire for transparency and responsiveness.

Interviewee 040, who also had previously worked for a repository, was also wary of repositories with an agenda. "It's tough because I think they add value but at the same time for the most part of what I've seen is they're actually marketing tools. It's like, 'Here we'll give you some free lessons and now buy this book of ours.' Or, 'This lesson is ... you can see how a teacher taught this lesson by buying this video of ours or buying this.'" Thus, transparency and knowing about repository actions can also have negative consequences not envisioned by Pirson and Malhotra (2011), particularly if stakeholders disagree with facets of the practices to which they've become privy.

Despite their interactions with repository staff, several of the interviewees were ambivalent or uncertain about the role of repositories with regard to curating, providing access, and preserving digital records of practice. Some had given it little thought. For example, when asked if repositories have trustworthy processes for curating and preserving data, 012 responded "Hopefully, probably. I mean, I think I would imagine so. … It doesn't cross my mind as a particularly big concern." 014 answered with uncertainty, "I don't think I understand enough about how things are saved to really be able to give any intelligent answer to that question." When asked whether she considers the trustworthiness of the repository when selecting data for reuse, 018 responded describing trustworthiness as a tacit consideration, "I suppose that's just something that is, you know, latent in your decision making, so I've never had to explicitly say that that's what I do, but that's definitely the case."

Trust in Repositories

Interview participants described several factors that contributed to their decisions about whether to trust a repository, including: data quality, co-production of data between the repository and the data producer, responsiveness of repository staff, and curation practices. Some of these factors align with Pirson and Malhotra's (2011) model of the depth of interaction or the closeness of affiliation.

Data Quality

Interviewees discussed indicators of trustworthiness, such as data quality, as markers that help them determine whether or not to trust a particular repository. Interviewee 002, who worked as a consultant for a digital repository, said that problems with data such as mislabeling can be indicators of bigger problems, and as a result she does not trust data with those kinds of errors, "If I see problems like that, I have no faith in the data and I won't dig into it." This is similar to what Lee et al. (2002) refer to as representational consistency or the extent to which data are presented in a professional manner, error free, and compatible over time. Representational consistency is a major indicator of credibility.

Interviewee 044 discussed the importance of curation processes and data quality in repository trustworthiness, "I think that in order for them to be valuable to teach education they have to have some sort of curation process. It has to be something that says that establishes a baseline of quality and I think we have that. We can't just let it be this wild thing where random videos are posted. It has to be at least a framework which gets us back to the idea that everything is curated with a particular thing in mind."

Co-production of Data between the Data Repository and Producer

Interviewee 008 trusted data from a particular repository because of her experiences working there and her trust in the data producer. "Part of the reason I work there is, I so completely believe in the product and [data producer's] work behind it." She went on to explain that she would be more likely to trust other repositories based on the recommendations or support of other researchers she knows and trusts. "Because I know and trust the validity, the reliability and the work that has been done to even

post those resources as valuable. So, it's almost that that's the piece that I look for, is I look for the experts out there to either have created their own or have said, 'Yeah, we really like [data producer's] work, so we're kind of putting a check next to his name and that's a good one.'"

Responsiveness of Staff

Among interviewees who interacted with repository staff, responsiveness emerged as another important trust factor. Previously, we quoted interviewee 037 who described in great detail above her experiences working with very responsive repository staff. Interviewee 002 explained that unresponsive repository staff and practices that lacked transparency were a barrier to developing trust with a repository, "Because I need to be able to trust the other decisions and things that they're telling me, and if I see a lot…I'm not looking for a typo and then writing something off, but if I see things that are obviously problematic, then I won't. And I think that the researchers, the original project managers need to be accessible. So, if you shoot off someone an email and you asked them about something… If they are not responsive, that's also a red flag for me."

Curation Processes

As noted in the previous sections, we found that knowledge of data curation processes was an underlying factor that intersected with other dimensions of repository trustworthiness. Interviewees who had experience working for a repository or serving on an advisory board, such as 037, 040, and 008, explained that their understanding of repository processes helped them to trust the repository and its data. "I've been pretty close to their process. I don't know about everything…I trust things that have – I'm not so convinced by [repository] or [repository], whatever it is. I don't know about them. I'm not sure who they're videotaping or why they're videotaping or how they're getting access to classrooms or what they're doing in terms of archiving their data. I think I feel more informed by [repository] or [organization], and so I'm comfortable" (037).

Interviewee 040 said that repositories "should be transparent." However, he also explained that his knowledge of repository processes was the result of participation on advisory boards rather than from information freely available to the general public, "The problem is – luckily we get invited to a lot of advisory boards because of this, because we have this video collection. We're working, we're on the advisory board for some other thing."

When discussing what makes her think that a particular repository is trustworthy, interviewee 026 explained that she believed the repository's processes for curating data were trustworthy because, "I know the people. I think the same thing of [repository], that they have a reputation. They're well-established and they care about things like that." She also trusted that the repository had the necessary permissions and rights to make video records of practice available, "For the [repository], I knew that whatever they put up, they had the right to put it up and make it available."

Interviewee 010 said that she trusts a particular repository because of how well they curated their collections, "one of the reasons I rely on [repository] so much is because I, they often do, they curate their videos, and I often appreciate their curation…I think they do a pretty good job of selecting the things that they put up there." She continued, "there's definitely things that I wouldn't show but in general, I trust most of their videos." Despite her overall trust for this particular repository, she was careful to note that she does not trust the repository's video descriptions, "I often don't trust those because there isn't I don't think, a common vocabulary for talking about teaching. Even if there is sort of kind of summary of what the video is, I don't often believe that that's going to tell me what I need to know. So, I'll often end up watching it anyway."

DISCUSSION

Pirson and Malhotra's (2011) model posited that stakeholder trust in repositories would align with two dimensions: intensity of data reuse and the closeness of the relationship or affiliation with the repository. We found evidence of both of these dimensions among our survey respondents and interviewees, but found that the closeness of the relationship was a stronger factor with regard to trust in digital qualitative data repositories containing video records of practice.

We found that our survey respondents and interviewees fell across the spectrum that Pirson and Malhotra (2011) described (i.e. depth of interaction and locus of relationship). 23.8% of survey respondents said that they are frequent data reusers, and 13.9% said that they have interacted with repository staff, both indicators of deeper interaction. Among interviewees, those who have worked at or consulted for a repository tended to have had the deepest interactions with repositories.

We also found strong evidence that a relationship or affiliation with a repository influences trust. The interviewees had a variety of types of relationships ranging from knowing a repository's reputation (minimal external relationship), to attending a workshop, to interacting with staff, to serving on an advisory board, being a consultant, or working as staff (intense internal relationship). Interviewees who spoke most strongly about trusting the repository had a stronger affiliation (e.g., interviewees 002, 004, 008, 010, 037, 040, 044). Furthermore, each of these individuals had some type of insider knowledge that gave them insight into the work of the repository and in particular its curation practices.

Factors such as data quality, the co-production of data between the data repository and producer, and staff responsiveness influenced trust in digital repositories. These factors, however, were confounded by curation practices. Knowledge of curation practices was raised alone as well as in conjunction with data quality and data repository / data producer actions. Thus, we identified knowledge of curation processes as the strongest link to repository trust. Our study found that trust in data extended to the metadata, and that the quality and care of the curation processes influenced trust in the repository. Previous research also demonstrates that data quality is a signal of trust in the data (Donaldson & Conway, 2015; Faniel & Jacobsen, 2010).

We identified some evidence that interactions between data reusers and repository staff in particular influenced their attitudes about, and trust in, digital repositories. Although reusers who interacted with repository staff were in the minority, these interactions were generally productive and those reusers tended to have positive attitudes toward repositories. No matter where data reusers fell along the shallow/deep encounters and internal/external affiliation with the repository, their interactions with repository staff often indicated areas where they experienced vulnerability. Overall, our interviewees appeared to have positive expectations when they asked questions, assuming that their questions would be answered (Pirson & Malhotra, 2011), indicating some measure of trust. Furthermore, reusers who interacted with repository staff only when they encountered a problem reported positive attitudes about those repositories when staff members were responsive, repository processes transparent, and the result was access to high quality data. This aligns with Pirson and Malhotra's research findings on transparency.

Research on trust in digital preservation suggests that repository users develop trust over time through repeated interactions (e.g. Kelton, Fleischmann, & Wallace, 2008; Yoon, 2016), and this is consistent with our findings. We found that data reusers value responsiveness and transparency of curation processes. Problems with data or other repository issues, if handled quickly and clearly, do not necessarily result in negative attitudes about repositories or a loss of trust.

Implications for Repositories

These findings suggest several implications for qualitative data repositories. Depth of interaction with repositories was correlated with trust, and the deep interactions with repositories that our interviewees reported were most frequently through official programs or channels (e.g. workshops, and consulting or advisory roles). This suggests that repositories wishing to cultivate relationships with users may want to consider additional ways to create opportunities for formal interactions and/or participation with members of their user communities over an extended period of time. This approach has been discussed more extensively in the archival research regarding participatory archives (e.g. Huvila, 2008). Our findings also suggest that transparency of repository practices, and especially data curation practices, are important for trust. This is consistent with findings from other studies (e.g. Yakel et al., 2013), and suggests that repositories should be transparent about the decisions underlying data curation and repository management practices whenever possible.

CONCLUSION

For data reusers, trust is essential at the data and the repository level. By adapting Pirson and Malhotra's (2011) model of depth of interaction with data or staff and closeness of the relationship to digital repositories with research data, we have found evidence demonstrating the importance of closeness of the relationship. The closeness is particularly important when it includes information about curation processes. This adds to our knowledge of trust in data and repositories, and provides support for Pirson and Malhotra's (2011) model for assessing that trust.

ACKNOWLEDGMENTS

REFERENCES

Akmon, D., Zimmerman, A., Daniels, M., & Hedstrom, M. (2011). The Application of Archival Concepts to a Data-Intensive Environment: Working with Scientists to Understand Data Management and Preservation Needs. *Archival Science*, *11*(3–4), 329–348. https://doi.org/10.1007/s10502-011-9151-4

Bak, G. (2016). Trusted by whom? TDRs, standards culture and the nature of trust. *Archival Science*, *16*(4), 373–402. https://doi.org/10.1007/s10502-015-9257-1

Bass, H., Usiskin, Z., Burrill, G., National Research Council (U.S.), Mathematical Sciences Education Board, & United States National Commission on Mathematics Instruction (Eds.). (2002). *Studying Classroom Teaching as a Medium for Professional Development Proceedings of a U.S.-Japan Workshop*. Washington, DC: National Academy Press. Retrieved from http://www.nap.edu/catalog/10289.html

Bishop, L. (2012). Using Archived Qualitative Data for Teaching: Practical and Ethical Considerations. *International Journal of Social Research Methodology*, *15*(4), 341–350. https://doi.org/10.1080/13645579.2012.688335

Burleigh, J. C., & Peterson, H. W. (1967). Videotapes in Teacher Education. *The Elementary School Journal*, *68*(1), 35–38.

Chowdhury, G. (2010). From Digital Libraries to Digital Preservation Research: The Importance of Users and Context. *Journal of Documentation*, *66*(2), 207–223. https://doi.org/10.1108/00220411011023625

Corti, L. (2000). Progress and Problems of Preserving and Providing Access to Qualitative Data for Social Research—The International Picture of an Emerging Culture. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *1*(3). Retrieved from http://www.qualitative-research.net/index.php/fqs/article/view/1019

Corti, L. (2011). The European Landscape of Qualitative Social Research Archives: Methodological and Practical Issues. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *12*(3). Retrieved from http://www.qualitative-research.net/index.php/fqs/article/view/1746

Derry, S. J., Pea, R. D., Barron, B., Engle, R. A., Erickson, F., Goldman, R., … Sherin, B. L. (2010). Conducting Video Research in the Learning Sciences: Guidance on Selection, Analysis, Technology, and Ethics. *Journal of the Learning Sciences*, *19*(1), 3–53. https://doi.org/10.1080/10508400903452884

Donaldson, D. R., & Conway, P. (2015). User Conceptions of Trustworthiness for Digital Archival Documents. *Journal of the Association for Information Science and Technology*, n/a-n/a. https://doi.org/10.1002/asi.23330

Donnelly, M., & North, R. (2011). The Milieu and the MESSAGE: Talking to Researchers about Data Curation Issues in a Large and Diverse e-Science Project. *International Journal of Digital Curation*, *6*(1), 32–44. https://doi.org/10.2218/ijdc.v6i1.170

Faniel, I., & Jacobsen, T. (2010). Reusing Scientific Data: How Earthquake Engineering Researchers Assess the Reusability of Colleagues' Data. *Computer Supported Cooperative Work (CSCW)*, *19*(3), 355–375. https://doi.org/10.1007/s10606-010-9117-8

Frank, R. D., Suzuka, K., & Yakel, E. (2016). Examining the Reuse of Qualitative Research Data: Digital Video in Education. In *Archiving Conference* (Vol. 2016, pp. 146–151). Washington, DC: Society for Imaging Science and Technology. https://doi.org/10.2352/issn.2168-3204.2016.1.0.146

Gaudin, C., & Chaliès, S. (2015). Video Viewing in Teacher Education and Professional Development: A Literature Review. *Educational Research Review*, *16*, 41–67. https://doi.org/10.1016/j.edurev.2015.06.001

Giarlo, M. (2013). Academic Libraries as Data Quality Hubs. *Journal of Librarianship and Scholarly Communication*, *1*(3). https://doi.org/10.7710/2162-3309.1059

Gonzalez, G., Deal, J. T., & Skultety, L. (2016). Facilitating Teacher Learning When Using Different Representations of Practice. *Journal of Teacher Education*. https://doi.org/10.1177/0022487116669573

Hadfield, M., & Haw, K. (2012). Video: Modalities and Methodologies. *International Journal of Research & Method in Education*, *35*(3), 311–324. https://doi.org/10.1080/1743727X.2012.717434

Holsti, O. R. (1969). *Content Analysis for the Social Sciences and Humanities*. Reading, MA: Addison-Wesley Pub. Co.

Huvila, I. (2008). Participatory Archive: Towards Decentralised Curation, Radical User Orientation, and Broader Contextualisation of Records Management. *Archival Science*, *8*(1), 15–36. https://doi.org/10.1007/s10502-008-9071-0

Karasti, H., & Baker, K. (2008). Digital Data Practices and Long Term Ecological Research Program Growing Global. *The International Journal of Digital Curation*, *3*(2), 42–58. https://doi.org/10.2218/ijdc.v3i2.57

Kelton, K., Fleischmann, K. R., & Wallace, W. A. (2008). Trust in Digital Information. *Journal of the American Society for Information Science and Technology*, *59*(3), 363–374. https://doi.org/10.1002/asi.20722

Lee, Y. W., Strong, D. M., Kahn, B. K., & Wang, R. Y. (2002). AIMQ: A Methodology for Information Quality Assessment. *Information & Management*, *40*(2), 133–146. https://doi.org/10.1016/S0378-7206(02)00043-5

Margaryan, A., & Littlejohn, A. (2008). Repositories and Communities at Cross-Purposes: Issues in Sharing and Reuse of Digital Learning Resources. *Journal of Computer Assisted Learning*, *24*(4), 333–347. https://doi.org/10.1111/j.1365-2729.2007.00267.x

Mauthner, N. S., & Parry, O. (2009). Qualitative Data Preservation and Sharing in the Social Sciences: On Whose Philosophical Terms? *Australian Journal of Social Issues*, *44*(3), 291–307,238.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. Thousand Oaks: Sage Publications.

Pea, R. D. (2006). Video-as-Data and Digital Video Manipulation Techniques for Transforming Learning Sciences Research, Education, and Other Cultural Practices. In J. Weiss, A. P. J. Nolan, J. Hunsinger, & P. P. Trifonas (Eds.), *The International Handbook of Virtual Learning Environments* (pp. 1321–1393). Springer Netherlands. Retrieved from http://link.springer.com/chapter/10.1007/978-1-4020-3803-7_55

Penn-Edwards, S. (2012). Human Factors Affecting the Use of Video Recording Methodology in Qualitative Research. *International Journal of Multiple Research Approaches*, *6*(2), 150–159. https://doi.org/10.5172/mra.2012.6.2.150

Pirson, M., & Malhotra, D. (2011). Foundations of Organizational Trust: What Matters to Different Stakeholders? *Organization Science*, *22*(4), 1087–1104. https://doi.org/10.1287/orsc.1100.0581

Poole, A. H. (2015). How has your science data grown? Digital curation and the human factor: a critical literature review. *Archival Science*, *15*(2), 101–139. https://doi.org/10.1007/s10502-014-9236-y

Pryor, G., & Donnelly, M. (2009). Skilling Up to Do Data: Whose Role, Whose Responsibility, Whose Career? *International Journal of Digital Curation*, *4*(2), 158–170. https://doi.org/10.2218/ijdc.v4i2.105

Rook, M. M., & McDonald, S. P. (2012). Digital Records of Practice: A Literature Review of Video Analysis in Teacher Practice (Vol. 2012, pp. 1441–1446). Presented at the Society for Information

Technology & Teacher Education International Conference. Retrieved from http://www.editlib.org/p/39784/

Russell, R., & Day, M. (2010). Institutional Repository Interaction With Research Users: A Review of Current Practice. *New Review of Academic Librarianship*, *16*(sup1), 116–131. https://doi.org/10.1080/13614533.2010.509996

Saldaña, J. (2015). *The Coding Manual for Qualitative Researchers* (3rd edition). Thousand Oaks, CA: Sage Publications.

Schofield, P. N., Bubela, T., Weaver, T., Portilla, L., Brown, S. D., Hancock, J. M., … Rosenthal, N. (2009). Post-Publication Sharing of Data and Tools. *Nature*, *461*(7261), 171–173. https://doi.org/10.1038/461171a

Shelton, C., Archambault, L., Warren, A., Ahmad, O., & Nevárez, V. (2016). Examining Preservice Teachers' Beliefs, Experiences, and Expectations Regarding the Use of Video in Education. In *Proceedings of Society for Information Technology & Teacher Education International Conference 2016* (pp. 465–468). Chesapeake, VA: Association for the Advancement of Computing in Education (AACE). Retrieved from http://www.learntechlib.org/p/171717

Stigler, J. W., Gallimore, R., & Hiebert, J. (2000). Using Video Surveys to Compare Classrooms and Teaching Across Cultures: Examples and Lessons From the TIMSS Video Studies. *Educational Psychologist*, *35*(2), 87–100. https://doi.org/10.1207/S15326985EP3502_3

van Es, E., & Sherin, M. (2002). Learning to Notice: Scaffolding New Teachers' Interpretations of Classroom Interactions. *Journal of Technology and Teacher Education*, *10*(4), 571–596.

Yakel, E., Faniel, I., Kriesberg, A., & Yoon, A. (2013). Trust in Digital Repositories. *International Journal of Digital Curation*, *8*(1), 143–156. https://doi.org/10.2218/ijdc.v8i1.251

Yoon, A. (2014). End Users' Trust in Data Repositories: Definition and Influences on Trust Development. *Archival Science*, *14*(1), 17–34. https://doi.org/10.1007/s10502-013-9207-8

Yoon, A. (2016). Data Reusers' Trust Development. *Journal of the Association for Information Science and Technology*. https://doi.org/10.1002/asi.23730

Zhang, M., Lundeberg, M., Koehler, M. J., & Eberhardt, J. (2011). Understanding Affordances and Challenges of Three Types of Video for Teacher Professional Development. *Teaching and Teacher Education*, *27*(2), 454–462. https://doi.org/10.1016/j.tate.2010.09.015

Zimmerman, A. S. (2008). New Knowledge from Old Data The Role of Standards in the Sharing and Reuse of Ecological Data. *Science, Technology & Human Values*, *33*(5), 631–652. https://doi.org/10.1177/0162243907306704