IJDC | *Research Paper*

Tool Selection Among Qualitative Data Reusers

Rebecca D. Frank
Berlin School of Library and Information Science,
Humboldt University of Berlin

Kara Suzuka
College of Education,
University of Hawai'i at Mānoa

Eric Johnson
College of Literature, Science and the Arts,
University of Michigan

Elizabeth Yakel
School of Information
University of Michigan

Abstract

This paper explores the tension between the tools that data reusers in the field of education prefer to use when working with qualitative video data and the tools that repositories make available to data reusers. Findings from this mixed-methods study show that data reusers utilizing qualitative video data did not use repository-based tools. Rather, they valued common, widely available tools that were collaborative and easy to use.

Received 15 December 2019 ~ *Accepted* 22 June 2020

Correspondence should be addressed to Rebecca D. Frank, Unter den Linden 6, 10099 Berlin, Germany. Email: rebecca.frank@hu-berlin.de

An earlier version of this paper was presented at the 15th International Digital Curation Conference.

The *International Journal of Digital Curation* is an international journal committed to scholarly excellence and dedicated to the advancement of digital curation across a wide range of sectors. The IJDC is published by the University of Edinburgh on behalf of the Digital Curation Centre. ISSN: 1746-8256. URL: http://www.ijdc.net/

Introduction

Repositories seek to support data reuse through a variety of activities and services. One area of support is the provision of tools to engage with, activate, and analyze the data. In the sciences, repositories, such GenBank[1] provide both data and tools that can be used to analyze that data, while other repositories, such as MorphoBank[2], have recreated the scientific workflows to enable data reusers to engage with the data to replicate or conduct new analyses. Repositories, such as SEAD[3], also provide tools and workflows that support the management, sharing, and reuse of data across multiple disciplines. In the quantitative social sciences, repositories such as ICPSR[4] provide data in a generic format (e.g., .csv) as well as in other formats (e.g., SPSS, R, Stata) in order to support data reuse with a variety of tools. For qualitative researchers, especially those reusing text-based data, some repositories also make data available in generic formats (e.g., .txt, .pdf). As with quantitative social science repositories, this enables data reusers to use their tools of choice when working with the data.

In contrast, some repositories of qualitative multimedia data (e.g., audio recordings, video recordings, still images), such as the Teaching and Learning Exploratory (TLE)[5], often expect users to access, view, and/or analyze data using tools provided by the repository within their own digital environment. This is the case for a variety of reasons, both practical and legal/ethical. For example, qualitative multimedia data have relatively large file sizes (e.g., .mp4), which makes them difficult for data reusers to download and manage. Another factor that contributes to repository decisions surrounding tools for multimedia data is the protection of research participants/subjects depicted in those data. For example, in the field of education qualitative data repositories preserve and provide access to qualitative data that frequently depict teachers and students in classroom settings. In these cases, repositories create secure spaces for analysis, such as the enclaves at the Inter University Consortium for Political and Social Research (Data Enclaves, 2019).

In this paper we explore the tension between the tools that data reusers in the field of education prefer to use when accessing video records of practice and the ways that repositories make those data available to data reusers. In particular we ask:

1. What types of tools do data reusers employ when they reuse video records of practice?
2. Why do reusers select these tools?
3. What are the implications of this tool selection/preference for repositories?

Our findings indicate that (1) data reusers in education use multiple tools when working with video records of practice; (2) they prefer common, widely available tools; (3) their preferences are influenced by two factors: collaboration and ease of use; and (4)

1 GenBank: https://www.ncbi.nlm.nih.gov/genbank/
2 MorphoBank: https://morphobank.org/
3 Sustainable Environment/Actionable Data (SEAD): http://sead-data.net/
4 Inter-university Consortium for Political and Social Research (ICPSR): https://www.icpsr.umich.edu/icpsrweb/
5 The Teaching and Learning Exploratory (TLE): https://tle.soe.umich.edu/

these preferences are often at odds with the ways in which repositories make video records of practice available to users.

Literature Review

Video Records of Practice

In the education field, Bass et al. (2002) describe records of practice as "detailed documentation of teaching and learning" (Bass et al., 2002), which can be analog or digital, include video of actual classroom activities, audio, still images, lesson plans, assignments, student work, and seating charts of classes. These records document different perspectives of the practices surrounding teaching and learning in an educational environment. In this paper we focus on digital records of practice, specifically video records of classroom activities. These present challenges for both data reusers and repositories, particularly because they require additional software for different types of reuse.

Video records of practice have been used for decades in educational research, in teaching undergraduate and graduate students how to teach, and in professional development or continuing education for teachers (Burleigh and Peterson, 1967; Shelton, Archambault, Warren, Ahmad and Nevárez, 2016; Stigler, Gallimore and Hiebert, 2000). For example, in teacher education and professional development settings, video records of practice are used to help build professional vision and noticing skills (Penn-Edwards, 2012; Rook and McDonald, 2012; van Es and Sherin, 2002). In research, video records of practice capture the evidence needed to study teaching and learning in an authentic setting (Hadfield and Haw, 2012).

In spite of these benefits, video records of practice are also challenging to reuse. Pea (2006) noted a gap between the promise and usefulness of video records of practice: "The availability of such inexpensive videography equipment and promise of more complete records of complex phenomena than earlier methods has led many researchers to adopt video recording as a primary data collection method. Yet, there is a serious and persistent gap between such promise and the usefulness of video records." Rich and Hannafin (2009) noted that "the amount of effort required to reflect on captured video influences both the processes and outcomes of reflection". Others have documented the time and training required to edit videos (e.g. Collins, Cook-Cottone, Robinson and Sullivan, 2004). Zhang et al. (2011) discussed ongoing technology problems with video data storage and sharing.

The problems and challenges of tools are shared by other data reusers. For example, Davidson, Paulus, and Jackson (2016) argue that qualitative researchers require knowledge and skills of digital tools to conduct research. Referring to scientific data reuse, Borgman et al. (2016) have advocated for "software as a service" provided by repositories of collaborative science laboratories; however, this solution will not work in highly heterogeneous data reuse environments, such as video data reuse.

Collaboration

Research about data reuse has examined the unique needs of researchers who work together across institutions, across physical locations/distances, and across disciplines,

and has identified the need for collaboration as a key factor in facilitating data sharing and reuse (Beaulieu et al., 2016; Toye, Sheppard, and Chen, 2016). Indeed, Jirotka, et al. (2013) argued that the success of e-Science collaborations involving data sharing and reuse, "relies fundamentally upon collaboration and the tools and technologies developed to support that collaboration within local and across global communities of researchers."

Much of the research about collaborative tools for data sharing and reuse has focused on the development of specialized technologies to facilitate particular types of work and communication (e.g. Androulakis et al., 2009; Jirotka et al., 2013; Kandogan et al., 2015). In education, Rich and Hannafin (2009) have argued that while tools for working with VRPs proliferate, it is important to understand the affordances and drawbacks of tools in order to make an informed decision about which will be most effective. Rather than ask what functionalities new tools and systems should have, or which tools are optimal for specific purposes, in this paper we instead ask what tools data reusers select for their own work, and what factors influence their selection decisions.

Ease of Use

Researchers examining data reuse have identified ease of use as a factor in data reuse. Huang et al. (2012) identified ease of manipulation (along with interpretability, understandability, consistent representation, and value added) as one dimension of the larger concept of usefulness for data reuse. Of the five data quality constructs, usefulness ranked in middle of the arithmetic averages of the mean ratings of the individual dimensions loaded on the constructs.

This seems to indicate that while ease of use is a factor in information quality and data reuse, it may be a mediating factor. This might be a result of the fact that a substantial amount of data reuse requires internal (to the repository) or external technologies to actually use the data. The role that technology plays in the use of video was directly addressed by Ju and Albertson (2018). They found a correlation between perceived ease of use of the technology and users' intentions to use video digital libraries (Ju and Albertson, 2018). In this paper we examine how data reusers select tools when reusing VRPs.

Research Methods

This mixed methods study employed a survey and qualitative semi-structured interviews with video data reusers in education. We detected patterns of tool selection and usage in our sample of video records of practice reusers through the survey. Then, we delved more deeply into the rationale and motivations for tool selection in the interviews (Creswell, 2009).

This paper draws upon data from a larger study that aimed to understand qualitative data sharing and reuse practices among researchers and teacher-educators in the field of education. This study was reviewed and deemed exempt by the Institutional Review Board at the University of Michigan.

Survey

Population and sample

The population for our survey consisted of reusers of repositories holding video records of practice. We surveyed users of four data repositories:

- Accomplished Teaching, Learning, and Schools (ATLAS) case library[6];
- Everyday Mathematics Virtual Learning Community (VLC)[7] at the University of Chicago;
- Inter-university Consortium for Political and Social Research (ICPSR), specifically Measures of Effective Teaching (MET) Collection[8] at the University of Michigan; and
- Teaching and Learning Exploratory (TLE) at the University of Michigan School of Education[9].

The ATLAS case library is associated with the National Board for Professional Teaching Standards. The 1,300 curated videos of teachers in this collection were assessed as accomplished practice, using National Board Standards.

The VLC enables teachers to access and share resources (including video records of practice) that support the teaching of mathematics. The collection consists of 446 videos and 782 associated records of practice.

ICPSR is a data archive consisting of over 250,000 datasets in the social and behavioral sciences, including 21 specialized collections in education. One of the educational data collections is the MET collection, which features longitudinal quantitative and qualitative data (including classroom videos) from 741 teachers in 317 schools (2009-2010) and 2,086 teachers in 310 schools (2010-2011).

The TLE offers classroom videos and curated collections of clips from over 1,000 sessions representing all K-12 grade levels and diverse socio-economic settings. Interaction with the videos is through a variety of applications to facilitate research or teaching reuse.

Three of these repositories focused primarily on reuse for *teacher education* purposes (ATLAS, VLC, and TLE) and one focused on reuse for *research* (ICPSR).

Survey responses

The survey was administered to 872 individuals, with a response rate of 20.6 percent (180 respondents). The population for our survey consisted of reusers of repositories holding video records of practice. For our analyses, we focused on the 149 respondents who reported that they had reused VRPs for research and/or teaching, representing 82.78% of our total respondents.

These 149 respondents were then separated into two groups depending on how they identified their primary reuse aim: research or teaching. If they reported that they had reused VRPs for both, they were asked to specify their primary focus for reuse.

6 The National Board for Professional Teaching Standards Accomplished Teaching, Learning and Schools (ATLAS): http://www.nbpts.org/atlas

7 Everyday Mathematics Virtual Learning Community (VLC): http://vlc.cemseprojects.org/

8 Measures of Effective Teaching Longitudinal Database (MET): https://www.icpsr.umich.edu/icpsrweb/METLDB/

9 Teaching and Learning Exploratory (TLE): https://tle.soe.umich.edu

Table 1. Summary of Survey Respondents by Primary Data Reuse.

Primary Reuse	Number of Respondents (N = 149)	Percentage of Respondents
Research	65	43.62%
Teaching	84	56.37%
Total	149	100%

Survey questionnaire

The survey consisted of four sections: questions about data reuse for research, data reuse for teaching, data reuse for personal study, and general questions about data reuse and participant background information. While all survey respondents were presented with the general questions, they were asked screening questions that determined whether they would see the research, teaching, and/or personal study questions, which were presented only when relevant to their own experiences. Questions about tools and services that respondents encountered while reusing qualitative video data were located in the general questions category that all respondents saw, regardless of their reuse goals.

Survey questions were developed based on themes from the literature as well as themes that arose from initial interviews. The survey questionnaire was tested through a series of cognitive interviews as well as a pilot test. The addition of the personal study category of questions was a made as a result of this initial pilot test.

Survey data analysis

The survey focused on baseline descriptive questions to understand how video records of practice data reusers actually interacted with these data. For this paper we draw from the questions concerning what type of tools they used and encountered when reusing VRPs, as well as open-ended questions asking them to report on the specific tools that they have used, and to discuss any challenges that they faced.

Interviews

Participants

We conducted in-depth semi-structured interviews with 44 video data reusers in the field of education, focusing on researchers and teacher-educators. Interviewees were recruited through a combination of convenience and snowball sampling. We first consulted relevant research literature and conferences in order to identify and recruit participants who had experience with data reuse, and also asked interviewees to nominate additional interviewees. The analysis for this paper includes 42 of those interviews, focusing on participants whose data reuse activities involved research and/or teaching.

Data collection: In-depth semi-structured interviews

Interviews lasted approximately one hour and included questions about participants' data reuse practices, including their attitudes about data reuse, repository practices, challenges encountered during data reuse, and the tools that they used during their data reuse experiences.

The interview protocol was developed using themes identified from a review of the literature, as well as themes that arose during the first ten interviews, which served as a pilot test for the interview protocol. Interviews consisted of three sections: questions about data reuse for research, questions about data reuse for teaching, and general data

reuse questions. Interviewees were asked to identify whether their data reuse focused primarily on research or teaching, and were asked the questions that were relevant to their experiences. All interviewees were asked the general data reuse questions.

Interview data analysis

All interviews were audio recorded and transcribed for analysis.[10] The interview transcripts were analyzed using NVivo, a qualitative data analysis software package. Initial codes were developed from the literature as well as themes that emerged from the interviews themselves. This approach included deductive and inductive approaches and utilized descriptive, analytic, and thematic codes. This process allowed us to compare our data with existing themes from the literature (Miles and Huberman, 1994; Saladaña, 2015).

Through this process, we developed a code set that included codes about data reuse, interactions between data producers, reusers, and repository staff, repository practices, and tools. This paper focuses primarily on the codes relating to tools that interviewees encountered during their data reuse experience.

We coded the interview data in two groups: interviewees whose reuse focused primarily on research, and those focused principally on teacher education. For each group, two coders worked independently coding the same transcript, repeating the process until we reached an acceptable level of interrater agreement. Using Scott's pi, a statistic to measure interrater reliability for coding textual qualitative data, we achieved a score of 0.712 for the research focused interviews and 0.732 for the teacher education focused interviews (Holsti, 1969; Scott, 1955).

In addition, we performed second-level axial coding and analysis on the codes relating specifically to tool use. After compiling a list of every tool mentioned during the interviews, we sorted them into categories based on those used in the survey. We also coded each tool based on whether or not they were web-based, and whether or not they enabled collaborative work (e.g., whether more than one person could work on the same document at the same time). This secondary analysis was conducted by two members of the research team, and we did not calculate interrater reliability for this analysis.

Findings

Survey respondents were asked to select categories of tools, based on the functionality (e.g., Analysis, Editing, Multimedia Authoring, Presentation, and Sharing, and/or Commenting) they employed when reusing video data. Presentation tools, such as PowerPoint were the most frequently used by survey respondents whose primary reuse focused on research (38 out of 65 respondents) and teaching (54 out of 84 respondents). See Table 2 below.

10 The interview protocol, qualitative data analysis codeset, and attributes for the interviews are available via the Institutional Repository at the University of Michigan: https://doi.org/10.7302/Z28C9TGP

Table 1. Types of Tools Used by Survey Respondents.

Tool Category	Research (n = 65)	Teaching (n = 84)
Analysis	21	8
Editing	14	29
Multimedia Authoring	3	7
Presentation	38	54
Sharing and/or Commenting	13	30

Interview participants mentioned over 48 specific tools. Researchers mentioned 30 different tools, and teacher-educators mentioned 34. Among interviewees, those whose reuse focused on teaching reported using an average of 4.16 specific tools when working with video records of practice. The highest number of tools used by teacher-educators was ten, but the mode for tool usage was one. Researchers reported using an average of four different tools when working with video records of practice. The highest number of tools used by any researcher interviewed was eight, and the mode was four. While teacher-educators reported using slightly more tools overall, individual researchers more commonly used multiple tools in their work.

Despite the prevalence of interviewees who reported working with multiple tools, some described negative experiences learning and working with a large number of tools. For example, Researcher_02 described the need to learn additional tools as overwhelming,

> 'It's a huge cognitive load because every single time I join a project or pick up a project or help or consult on a project I have to learn a new system or set of systems too 'cause there are usually three different things you have to learn, you have to add them all to your computer. Then once you learn them or have enough starting skills to interface with them, then whenever you're moving back and forth between data sets, it's not just the cognitive load of switching over between data sets, it's the cognitive load of switching over between tools.'

As a collaborator on multiple projects, she had come to use many different analytic tools over time. She went on to say that she has declined collaborating on projects because it would require learning new analysis software: "I really struggle with that to a point where I've actually said no to two projects this year because I had no desire to learn new coding software. I just wouldn't do it." For this interviewee, decisions about data reuse were influenced by the tools that she would have to use limiting her ability to collaborate on research projects.

Using the tool categories that we developed for the survey, we found that among interviewees analysis tools were the most commonly mentioned category by researchers, while editing tools were most commonly mentioned by teacher-educators. In this paper, we focus on usage of tools within two categories: analysis (primarily researchers) and editing (primarily teacher-educators).

We found that tool selection within these two categories (i.e., analysis and editing) was driven primarily by two main factors: collaboration and ease of use. In the sections below, we focus on the ways that the need for collaboration drove selection of analysis tools, and the ways that the need for easy to use tools drove selection of editing tools.

Analysis: Striving for Collaboration

Collaboration was a strong theme among the interviewees. For both research and teaching, the disciplinary norm of working with others was a factor that influenced the tool selection by both researchers and teacher-educators when working with VRPs. Collaboration was prominent in interviewee discussions about tools for data analysis. Nine researchers reported purposefully selecting analysis tools with collaborative capabilities. They indicated wanting multiple people in different locations to access and work with the same data concurrently.

Among those interviewees, the need to allow multiple collaborators to access and code data from disparate locations was a key factor in tool selection. For example, one interviewee (Researcher_17) explained that Dedoose became an important tool because it facilitated collaboration with partners who were geographically distributed:

> 'As my work has become more and more collaborative, I've started using Dedoose more … my collaborators are not even in the same state or in the same continent as me and so Dedoose is great because we can all look at and do some data analysis in the same group.'

For this researcher, the need to collaborate shifted tool selection toward a platform that allowed multiple people to work with the data simultaneously from different locations.

Researcher_41 explained that the collaborative aspects of ATLAS.ti made it easier for her team to coordinate their analysis work:

> '[W]e used a program called ATLAS.ti. We used that to tag and code student writing, their VMC analytics conversion. That was a nice way for us to organize our data between coders. Where we didn't have to write all of the comparison formulas that you would in Excel files.'

For this reuser, the collaborative affordances of ATLAS.ti were an improvement over her team's previous process, which involved doing the work by hand: "But we predominantly use(d) Excel and held in-person coding meetings and done comparisons by hand, at least early on that's what we were doing." Similarly, Researcher_40 also described a research process that relied on Dedoose to facilitate intercoder reliability. He explained that his team uses it partly because of its popularity among his research community, and partly because of the video support provided by Dedoose.

Another approach to collaborative analysis, and intercoder reliability in particular, was described by Researcher_12:

> 'In terms of the coding itself, I'm actually amazed at how much you can do with things like Google Spreadsheets now. So, for instance, when we were calculating interrater agreements, we could just set up a spreadsheet with five versions of that transcript, have people code on there, and then just write formulas to compare across the five sheets, and so we could actually calculate reliability that way. So that was really useful, versus something like R, you could do it, but it would be a lot more difficult to write the code. So that's … I'm impressed, especially now that I feel like most projects, most larger projects are these collaborations and often

> across institutions. So anything like the Google software that can help you collaborate across sites is extremely helpful.'

For this researcher, the ability to easily collaborate across institutions with a widely-available web-based tool that did not require coding skills was an important factor in selecting Google Sheets as a data analysis tool. When asked if he had used tools that were designed specifically for working with video, he said that even though there may be better tools for his work, the amount of work required to find and learn how to use them was a deterrent, "I mean, there are probably tools to much better do what I want to do, but I haven't invested the time to learn about them or how I would use them."

Researcher_13 also mentioned Google Sheets as an analysis tool, and explained that while this has been the primary analysis tool, they would explore other options if they could find something that would be useful for their work:

> 'Well, the main thing for us is just in making sure the entire team has access to all of the data and so when the team over there gave us access, the first thing we did was upload it to a secure server at [institution]. Not exactly software, but that's been really important for us. Mainly from there once the videos are transcribed, a lot of it has been done on Google Docs using their spreadsheet … That's just the primary tool. Next semester once we kinda switch into qualitative analysis, we have a video observation lab on campus in our Education building. I don't know what video analysis software they use but we're gonna also start using that if we think it's gonna be useful for us.'

For the researchers who discussed using collaborative analysis tools, the ability to work in the same document or project was important. A particularly important part of the research process for this collaborative work was in the early stages of analysis when they were coding for interrater reliability.

Video Editing: Ease of Use

We found that interviewees valued convenience in their video editing tools. Several researchers and teacher-educators explained that they tended to use whatever program was installed and readily available on their computers rather than consciously selecting a particular tool for video editing, and also that they switched tools only when they needed additional capabilities or when a new tool would make their work easier. For example, Teacher-Educator_33 said that he used more than one editing tool, switching to iMovie when QuickTime would not meet his needs, "QuickTime I think is the one that's generally the one that's the default. I've used iMovie to make my own and then put them up." Interviewees also chose editing and analysis tools to use with video records of practice based on their disciplinary research collaborations or their teaching partners' ability to access and use the software. Ease of use was a major criterion for tool selection, and interviewees made decisions based on their own needs as well as those of their collaborators.

Teacher-Educator_36 explained that her needs tend to be pretty basic, and that she asks for help in working with video records of practice if she needs to do anything else, "I don't do a lot. If I really need to do some fancy clipping and stuff I ask my wife to help me with iMovie or stuff like that because I don't do that stuff very much."

Teacher-Educator_43 said that he switched to a more expensive tool he uses when editing video records of practice, and that the ease of use and time saving justified the cost,

> '[W]e capture it in HD but then we convert it to a lower resolution file. There's a program called Turbo.264 HD. It's super easy to use. You can basically take a file and you just drag and drop. It's got presets so I can convert it to iPod quality, an iPhone quality, an Apple TV quality. You just drop it in, you put the setting and it runs. It comes with a USB dongle that has a processor built in to help speed up the process. Before that I was trying to use iMovie or something like that to do the conversion and it was just painfully slow and difficult … That little program itself is worth every penny.'

Teacher-Educator_10 said that the tool she most often uses for video editing is QuickTime, "All my videos are in QuickTime, and so I'm constantly clipping and saving stuff in QuickTime." (Teacher-Educator_10). Similarly, Teacher-Educator_14 said that she typically uses QuickTime, but will switch to another tool such as VLC player depending on which will most easily enable the type of editing that she needs to do, "Mostly, really QuickTime. Sometimes I use a VLC player if QuickTime doesn't, for whatever reason, play the player, play it. Or if I need to up the video, up the audio, I can more easily do that through the VLC player." Teacher-Educator_32 also said that although she has worked with VLC, QuickTime was preferred by her collaborators, who were not particularly savvy about technology, "We all have Macs. I assume everybody else is using QuickTime. The people I work with are not super tech adept so I doubt they would even know that there is a thing called VLC."

Indeed, QuickTime was a very popular choice among our interviewees. In addition to Teacher-Educators 10 and 14 above, Teacher-Educators 23, 24, 25, 26, 30, 32, 33, and 38 also said that they use QuickTime when editing and viewing video records of practice for their teaching because it is the most convenient tool for them, "I think usually the default is QuickTime" (Teacher-Educator_26). Teacher-Educator_25 explained that while she worked with more specialized tools such as StudioCode for research, ease of use drove her decision to use QuickTime when working with video records of practice for her teaching, "But like I said, depending on like for research, it's been different. We've used it by putting it in StudioCode for things like that."

Although multimedia editing tools were more commonly discussed by teacher-educators, researchers also reported selecting editing tools that were easy to work with. Researchers 17, 18, 39, and 40 all said that they had used Windows Media Player or QuickTime, and that they tended to use these particular programs because it was often the default video player on their computers, "Usually Windows Media Player, something like that. Whatever is on the computer. QuickTime. If you use a Mac, usually use QuickTime" (Researcher_40).

Among our interview participants, ease of use was an important factor in the selection of editing tools. Interviewees tended to describe their tool selection process as one of switching to something new only when it was easier to use or provided additional features that they needed – or simply using the default programs that were already installed on their devices. In fact, several described their work not in terms of tool selection, but in terms of relying on default programs that were already installed on their devices such as QuickTime.

Implications for Repositories

Although several of the repositories used by our participants offered tools for working with their data, participants primarily commented on easily accessible, ubiquitous, or cloud-based tools not provided by repositories (e.g., QuickTime, Excel, Google Sheets, Dedoose). Reusers of video records of practice preferred selecting their own tools for their tasks. The number of tools used is also notable, particularly among researchers who routinely used four separate tools in reusing video. This implies that researchers are investing substantial time to learn tools and that there are coordination costs to collaborating with others in working with the video records of practice (e.g. Onal Vural, Dahlander and George, 2013).

Tool selection was influenced by the primary reuse goal (i.e., research or teaching). However, we also found that researchers valued tools that were collaborative and easy to use. The tools most frequently discussed by our interviewees tended to be widely popular, inexpensive, with features that were stable over time, and that were not domain-specific.

A notable example was the use of Google Sheets for coding video data. Rather than using a purpose-built analysis tool (e.g., StudioCode), or working within a repository environment (e.g., TLE), participants described a data analysis process whereby they would create a simple spreadsheet with a column to track the timestamp in the video and other columns for analysis fields such as codes and/or transcription of the video. This approach allowed participants to collaborate with others, regardless of their different computing environments, without requiring a significant investment of time and resources to obtain and learn new tools for working with video records of practice. The web-based tool allowed them to easily share their work with collaborators.

In general, we found that video records of practice data reusers did not use repository-based tools. Scholars, such as Kethers, Treloar, and Wu (2016) and Wolski et al. (2017), have argued for developing a centralized repository of tools to facilitate the data lifecycle, particularly data reuse. In the field of education, there are strong incentives for repositories to want data reusers to work within the repository environment (e.g., data enclaves to protect the privacy of children featured in video records of practice). However, our findings indicate that data reusers prefer to use their own tools, and in fact very few reported using the tools provided by repositories for working with video records of practice.

In this paper, we argue that in order to meet the needs of data reusers, data repositories should consider incorporating more ubiquitous tools, that are likely to be well-known and supported, rather than purpose-built tools that require users to learn new technologies in order to reuse data. In this way, they can accomplish their own data security goals while also meeting the needs of their users.

Acknowledgements

This project was made possible in part by the Institute of Museum and Library Services grant # LG-06-14-0122-14. We also acknowledge the University of Michigan Undergraduate Research Opportunity Program (UROP), which supported one of the authors.

References

Androulakis, S., Buckle, A. M., Atkinson, I., Groenewegen, D., Nicholas, N., Treloar, A., & Beitz, A. (2009). ARCHER – e-Research tools for research data management. *International Journal of Digital Curation, 4*(1), 22–33. doi:10.2218/ijdc.v4i1.75

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. Retrieved from http://www.nap.edu/catalog/10289.html

Beaulieu, S. E., Fox, P. A., Di Stefano, M., Maffei, A., West, P., Hare, J. A., & Fogarty, M. (2016). Toward cyberinfrastructure to facilitate collaboration and reproducibility for marine integrated ecosystem assessments. *Earth Science Informatics.* doi:10.1007/s12145-016-0280-4

Borgman, C. L., Golshan, M. S., Sands, A. E., Wallis, J. C., Cummings, R. L., Darch, P., & Randies, B. M. (2016). Data management in the long tail: Science, software, and service. *International Journal of Digital Curation, 11*(1). doi:10.2218/ijdc.v11i1.428

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

Collins, J. L., Cook-Cottone, C. P., Robinson, J. S., & Sullivan, R. R. (2004). Technology and new directions in professional development: Applications of digital video, peer review, and self-reflection. *Journal of Educational Technology Systems, 33*(2), 131–146. doi:10.2190/G9E9-LCQ1-H50X-4N23

Creswell, J. W. (2009). Research design: Qualitative, quantitative, and mixed methods approaches. Los Angeles: Sage.

Data Enclaves. (2019). Retrieved December 13, 2019, from Inter-university Consortium for Political and Social Research website: https://www.icpsr.umich.edu/icpsrweb/content/ICPSR/access/restricted/enclave.html

Davidson, J., Paulus, T., & Jackson, K. (2016). Speculating on the future of digital tools for qualitative research. *Qualitative Inquiry.* doi:10.1177/1077800415622505

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

Holsti, O. R. (1969). Content analysis for the social sciences and humanities. Retrieved from http://mirlyn.lib.umich.edu/Record/000000426

Huang, H., Stvilia, B., Jörgensen, C., & Bass, H. W. (2012). Prioritization of data quality dimensions and skills requirements in genome annotation work. *Journal of the American Society for Information Science and Technology, 63*(1), 195–207. doi:10.1002/asi.21652

Jirotka, M., Lee, C. P., & Olson, G. M. (2013). Supporting scientific collaboration: Methods, tools and concepts. *Computer Supported Cooperative Work (CSCW), 22*(4–6), 667–715. doi:10.1007/s10606-012-9184-0

Ju, B., & Albertson, D. (2018). Exploring factors influencing acceptance and use of video digital libraries. *Information Research, 23*(2). Retrieved from http://InformationR.net/ir/23-2/paper789.html

Kandogan, E., Roth, M., Schwarz, P., Hui, J., Terrizzano, I., Christodoulakis, C., & Miller, R. J. (2015). LabBook: Metadata-driven social collaborative data analysis. *2015 IEEE International Conference on Big Data (Big Data),* 431–440. doi:10.1109/BigData.2015.7363784

Kethers, S., Treloar, A., & Wu, M. (2016). Building tools to facilitate data reuse. *International Journal of Digital Curation, 11*(2). doi:10.2218/ijdc.v11i2.409

Miles, M. B., & Huberman, A. M. (1994). Qualitative data analysis: An expanded sourcebook. Thousand Oaks: Sage Publications.

Onal Vural, M., Dahlander, L., & George, G. (2013). Collaborative benefits and coordination costs: Learning and capability development in science: Collaborative benefits and coordination costs. *Strategic Entrepreneurship Journal, 7*(2), 122–137. doi:10.1002/sej.1154

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). 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. doi:10.5172/mra.2012.6.2.150

Rich, P. J., & Hannafin, M. (2009). Video annotation tools: Technologies to scaffold, structure, and transform teacher reflection. *Journal of Teacher Education, 60*(1), 52–67. doi:10.1177/0022487108328486

Rook, M. M., & McDonald, S. P. (2012). Digital records of practice: A literature review of video analysis in teacher practice. Proceedings of the Society for Information Technology & Teacher Education International Conference, 2012, 1441–1446. Retrieved from http://www.editlib.org/p/39784/

Saldaña, J. (2015). Thinking qualitatively: Methods of mind. Thousand Oaks, California: SAGE.

Scott, W. A. (1955). Reliability of content analysis: The case of nominal scale coding. *Public Opinion Quarterly, 19*(3), 321. doi:10.1086/266577

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. Proceedings of Society for Information Technology & Teacher Education International Conference 2016, 465–468. 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. doi:10.1207/S15326985EP3502_3

Toye, G., Sheppard, S., & Chen, H. L. (2016). Data sharing and reuse within the Academic Pathways Study. *Advances in Engineering Education, 5*(2), 19.

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.

Wolski, M., Howard, L., & Richardson, J. (2017). The importance of tools in the data lifecycle. *Digital Library Perspectives, 33*(3), 235–252. doi:10.1108/DLP-11-2016-0042

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. doi:10.1016/j.tate.2010.09.015