

Complementary Data as Metadata: Building Context for the Reuse of Video Records of Practice

Allison Rae Bobyak Tyler
University of Michigan School of
Information, United States of America

Kara Suzuka
University of Hawaii at Mānoa, United
States of America

Elizabeth Yakel
University of Michigan School of
Information, United States of America

Abstract

Data reuse is often dependent on context external to the data. At times, this context is actually additional data that helps data reusers better assess and/or understand the target data upon which they are focused. We refer to these data as complementary data and define these as data external to the target data which could be used as evidence in their own right. In this paper, we specifically we focus on video records of practice in education. Records of practice are a type of data that more broadly document events surrounding teaching and learning. Video records of practice are an interesting case of data reuse as they can be extensive (e.g., days or weeks of video of a classroom), result in large file sizes, and require both metadata and other complementary data in order for reusers to understand the events depicted in the video. Through our mixed methods study, consisting of a survey of data reusers in 4 repositories and 44 in-depth interviews, we identified the types of complementary data that assist reusers of video records of practice for either teaching and/or research. While there were similarities in the types of complementary data identified as important to have when reusing VROP, the rationales and motivations for seeking out particular complementary data differed depending on whether the intended use was for teaching or research. While metadata is an important and valuable means of describing data for reuse, data's meaning is often constructed through comparison, verification, or elucidation in reference to other data.

Submitted 28 December 2019 ~ Accepted 19 February 2020

Correspondence should be addressed to Allison R.B. Tyler, 105 S State St, Ann Arbor, MI 48109. Email: arbtyler@umich.edu

This paper was presented at International Digital Curation Conference IDCC20, Dublin, 17-19 February 2020

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/>

Copyright rests with the authors. This work is released under a Creative Commons Attribution Licence, version 4.0. For details please see <https://creativecommons.org/licenses/by/4.0/>



Introduction

The classic definition of metadata is ‘data that describes data’ (Gilliland, 2016). Yet, when we think of metadata we most often think of descriptive metadata: short, brief words or phrases, highlighting salient characteristics of the data at hand. However, in order to meaningfully reuse of data over time, the definition proposed by the DESIRE Project is perhaps more apt: ‘Data associated with objects which relieves their potential users of having to have full advance knowledge of their existence and characteristics’ (Dempsey et al., 1997, p. 2).

In this paper, we discuss data reuse and data reusers’ needs for additional data to help them understand the data they intend to use. We refer to these data as complementary data and define these as data external to the target data which could be used as evidence in their own right. Specifically we focus on video records of practice (VROP) in education. Records of practice are a type of data in the field of education that more broadly document teaching and learning events. These data include lesson plans, student work, seating charts, and video of classroom activities. We are particularly interested in video records of practice because they can be extensive (e.g., days or weeks of video of a classroom), result in large files sizes, and require both metadata and other complementary data in order for reusers to understand the events depicted in the video.

In our mixed methods study, comprised of a survey and interviews, educational researchers and teacher-educators indicated the need for a number of different types of data, which we refer to as complementary data, to better use the video records of practice. As a result we pose the following research questions: 1) What types of complementary data are important during reuse? 2) What is the motivation for needing these data? and 3) What are the implications for repositories?

Literature Review

Context from Digital Objects to Research Data

Although more data are becoming available for reuse, data often require additional context to make that reuse meaningful. Initial research on what additional context was needed for reuse focused on the significant properties of digital objects. Hedstrom and Lee (2002) developed a model ‘to identify the significant properties of digital objects that affect their quality, functionality, and look-and-feel’ (p. 218). In a subsequent study focusing on word processing/outlining programs and computer games, Hedstrom, Lee, Olson and Lampe identified three types of critical contextual information: ‘information about the context in which the objects were originally created and used; information about the purpose and audience for the materials; and information about the original computing environment’ (2006, p. 187). In a related article, Lee (2011) described context more broadly through an explication of three categories of context. In the first type of context a digital object relies on its provenance or the physical or logical structure in which it is embedded. In the second type of context, characteristics of the creation of a digital object are key and in the third, the user, or in our case, the data reuser brings a context to the digital object or data. Neither Hedstrom et al. nor Lee are discussing data specifically. In a more recent article, Faniel, Frank, and Yakel (2019) identified the properties of data that reusers would like to have to interpret the data. These properties included both descriptive metadata elements (e.g., name of the data, data collection information) and actual complementary data (e.g., access to the original artefact or specimen for archaeologists and zoologists respectively).

Research examining when the context needed to understand data is actual data is sparse even though we have evidence that many disciplines require additional data in order to contextualize the target data and facilitate reuse. When examining mammography images, Hartswood et al. (2012) reported on radiologists' difficulties with interpretation. To overcome these difficulties, the radiologists needed to align the scans with additional information from the medical history, such as previous scans or test results. In their study of the Center for Embedded Networked Sensing (CENS), Wallis, Rolando, and Borgman (2013) found that sensor-collected data were frequently used to assess the validity of hand-collected data. However, the sensor data often were considered 'background data,' not 'primary data,' and therefore were neither reported in publications nor kept for future use. The authors argued that such background data can be valuable—even essential—to reusers of the shared primary data (for instrument calibrations, result comparisons, and validation) or for others doing related work for whom the "background data" might serve as "foreground data."

Video Records of Practice

The need for additional data is particularly important when using video records of practice. Hatch and Grossman (2009) noted the role of additional data, such as lesson plans, curriculum materials, and demographic data to enable researchers and teachers to fully make use of the video data. Andersson and Sørvik (2013) also pointed to the need for additional data, such as student work, curriculum materials, and other teaching materials, when reusing video records of practice. They argued that these additional data enabled a wider variety of future studies. These examples indicate the importance of additional data, separate from but related to the data targeted for reuse, for verification, ascertaining meaning, and making the most out of existing data for new research questions.

Methods

We report on a mixed methods study that employed a survey and qualitative semi-structured interviews. We used the survey data to identify data reusers' needs concerning complementary data, then analysed the interviews to examine their rationale and motivations. The survey was administered to data reusers at 4 repositories holding video records of practice. Three of these repositories primarily targeted teacher-educators (although one served both teachers and researchers to a certain extent) and one primarily focused on researchers. We administered the survey to 872 individuals. This translated into 180 completed surveys for a response rate of 20.64%. In this article, we report only on the 149 respondents indicating they reused VROPs for research and/or teaching. In addition, we conducted in-depth semi-structured interviews with 44 researchers and teacher-educators who reused video data. We first consulted relevant research literature and conference programs/proceedings to identify and recruit video data reusers; then we asked interviewees to nominate additional interviewees. The analysis for this paper includes 42 interviewees who reused VROPs for research and/or teaching. Questions centred on participants' reuse of VROPs, what information they required to meaningfully reuse VROPs, and the problems they identified in data reuse. The interview protocol was developed using themes identified from a review of the literature, as well as themes that arose during the initial interviews.

Findings

Overall, our interviewees articulated a need for complementary data to add perspective or make the video records of practice 'complete.' In the words of Interviewee_017, '[W]hen you are

looking at a video blind, like it's not a video you collected, the more information there is about it the better.' In this section, we delve more deeply into the type of complementary data desired and the reasons they are important in data reuse.

Survey respondents were asked whether 6 types of complementary data were important for their use of video records of practice. Interviewees also mentioned these 6 categories of data. Comparing the results from these two data collection methods, we found that interview and survey participants identified similar types of complementary data to use alongside the video records of practice whether they were using the video for research or teaching. There were differences in the intensity of the stated need, particularly among the interviewees. Table 1 shows the frequency (percentage) of complementary data desired by the survey and interview participants.

Table 1. Frequency (percentage) of complementary data use for research and teaching.

	Research		Teaching	
	Survey (n=65)	Interview (n=32)	Survey (n=84)	Interview (n=36)
Lesson Materials	55.4	46.9	72.6	47.2
Student Work	53.9	40.6	77.4	38.9
Transcripts	40.0	65.6	27.4	58.3
Lesson Plans	38.5	40.6	70.2	54.2
Seating Charts	10.8	15.6	11.9	2.8
Classroom Photos	20.0	—	21.4	—

Survey respondents could also write-in responses to a follow-on survey question about additional complementary data use. These respondents also indicated that student demographic data and test scores were useful. Interviewees additionally provided a much richer list of complementary data which were not asked about in the survey. These data included school and classroom demographic data, information about the teacher characteristics and demographics, and teacher diaries, journal entries, or reflections about the recorded classroom activities.

While there were similarities in the types of complementary data identified as important when reusing video records of practice, study participants' rationales and motivations differed depending on whether the intended use was for teaching or research. One interviewee summarized this when we asked, 'Is it also important to you to have things like copies of student work or copies of the particular math problem that they're working on in a video?' Interviewee_026 replied, '...[I]t's really dependent on how I'm using the video.' The interviews provided an opportunity to better understand the motivations and methodological need behind the use of different complementary data. In the next section we discuss our analysis of the interviews, focusing on the four most frequently identified and also co-occurring types of data: transcripts, class materials, lesson plans, and student work.

Transcripts

Transcripts of the videos helped researchers examine classroom events through multiple frameworks, especially when paired with student work or the lesson plans. Similarly, when used for teaching purposes, transcripts offered opportunities to show/see more than might be seen with video alone. According to one teacher-educator, transcripts served as 'a more powerful way of showing like what the kids were actually doing ... using this transcript. It was the first time that I've used a transcript, and I felt like the transcript actually was able to come alive' (Interviewee_003). The videos showed students working on assignments, but do not always

capture what it is the students are doing. The copies of the student work, when paired with transcripts of student-teacher or student-student conversations, helped both researchers and teacher-educators (and the pre-service teachers or actual teachers with whom they work) better understand what was happening in the video.

‘Having a transcript along with it, along with the video. For a few reasons 'cause then you can actually hear better when you're actually reading along, but you can also go back in the discussion and use evidence like ‘Oh, here's where I think she under...This child understood what's going on and this is where she lost it or whatever.’ You use their words.’ (Interviewee_015)

Class or Lesson Materials

Data reusers needed additional context for class activities from different perspectives to make sense of the video records of practice. This was provided through class or lesson materials which we defined as class handouts or information distributed by the teacher to further the aims of a given lesson. These included the documents given to students, such as worksheets, and information written on the whiteboard during the lesson. In terms of the specific context created by class materials, teacher-educators emphasized the importance of understanding the totality of class activities, the ability to gain insight into the teacher’s decisions, and the desire to use these materials as instructional aids when educating teachers.

The context of class activities in professional development and teacher education was paramount. Viewing the video was important, but Interviewee_005 notes that without the class exercise, making sense of the video was compromised:

‘...[I]t's useful particularly when you're using videos for teacher education purposes to have the tasks recorded someplace. So sometimes what can happen in the data set is that it's pretty hard to figure out the math that people were working on...In a particular video, or the recording starts after the task has been launched and you don't know what the wording of the task was...if kids are working on a worksheet having access to a scanned copy of even just the blank worksheet to contextualize what's going on, so things like that I think are quite useful.’

Teaching reuse often involved participants actually doing the math problem itself before viewing videos to reflect on the teaching and to understand the level of difficulty the students were encountering. Interviewee_001 argued:

‘[W]hen we have teachers or other observers watch video of a math lesson, we almost always tell them what the problem was that the kids are gonna be working on and to actually do the math and solve the problem, so that they have some sense, they get some idea of what space the kids are in and how difficult that might or might not be for them. And without doing that, you're very limited in how to interpret what you see happening with the kids.’

Researchers also described the need for class materials. For example, Interviewee_039 discussed the ability to triangulate data from the video with the class materials:

‘[H]aving scanned material from a given lesson is extremely useful for me, knowing exactly, for example, if students in the video were given a worksheet to work on, having copies of those is useful...I always collect, in addition to the video, I collect the materials that were developed for that lesson or that were

produced in the context of the lesson and I use those to triangulate the findings or to understand better what happened in the classroom.’

This idea of having the full context and being able to compare or triangulate the data with the video data was consistent across resuers for teaching or research purposes.

Lesson Plans

Similarly, lesson plans were important for providing context for what the teacher intended to happen in the classroom. In particular, lesson plans were used to determine how a class fit into the larger curriculum as well as to gain insight in teachers’ thought processes. According to one researcher,

‘Knowing what the students have done in the past, knowing what they're going to be doing in the future, knowing what the teacher's goals are for their lessons, makes it a lot easier to make sense of teacher moves. Because then you can see, oh the teacher's doing this because they did this the other day, or because the teacher's goal is this.’ (Interviewee_034)

Understanding how a given lesson fit into the entire curriculum was essential for both researchers and teachers. Knowledge of the larger curriculum ranged from a desire to know whether ‘the lesson is directly building off another lesson’ (Interviewee_025) to a need for a more longitudinal context to ‘get a better understanding of the types of things the kids have already done and thought about and talked about, and the direction where the teacher [i]s anticipating going’ (Interviewee_034). For teacher-educators, the lesson plans also were important for framing the lesson for viewers. For example, a teacher-educator explained, ‘One of the things that I think is always important is framing a lesson. Where is this lesson in the series of lessons or unit that they're doing? ...What's come before or what's come after or how does this fit in?’ (Interviewee_043)

For researchers, lesson plans were seen as part of a larger context and in many ways, a given lesson was only understood in terms of the larger arc of the curriculum—what Interviewee_011 referred to as the ‘trajectory’ of the lessons. One researcher noted,

‘Context like I always like to know if we're watching a segment or a lesson, where is that embedded, what is the larger picture. If it's a clip, I want to know it's part of what lesson where in the lesson. The same is for a lesson. I want to know what was the unit about, at what point of the unit was this lesson taught. Some contextual information on the video is very important. (Interviewee_039)

Lesson plans were also used to gain insight into teachers’ thought processes. Lesson plans could be the formal lesson planning documentation or teachers’ journal entries or reflections concerning the goals for a lesson. Interviewee_003, a researcher, argued that ‘the lesson plans are crucial.’ For her, the lesson plans defined the aims of the lesson but also defined what constituted a lesson. She said that lesson plans outlined the ‘Purposes, goals, what the teachers saw as a segment of instruction for example.’ Interestingly, her research team used the lesson plans in conjunction with the transcript: ‘So like, it can be confirmatory, ‘Oh, these were the goals that the teacher set for that day.’ ...as we do the transcripts we'll look again at the lesson plans, we go back to the lesson plans quite a bit.’

In terms of teacher education, the lesson plans were not only used to give insight into the thought processes of the teachers in the video, but also to help hone the thought processes of the pre-service teachers. Interviewee_020 described her instruction of student teachers:

‘So I mean we've also kind of orchestrated these activities so that they look at the video, and then we have them read what the teachers and the researchers were thinking about, because we don't want them to be tainted when they're thinking on their own about what's happening, and we'd kinda like them to contrast what they thought initially with what they thought later, after they saw the video and read what the teachers were thinking. Having that really good text behind it, I think is another factor...having those lesson plans for them to look at. It makes a combination and you can really leverage the whole program, not just one piece.’

These examples show how lesson plans are important to both researchers and teachers: providing insight into the larger trajectory of the curriculum as well as how teachers think about helping learners move forward on that trajectory. Lesson plans are often not used in isolation; transcripts, teacher reflections, and the video records of practice help complete the evidential or instructional picture.

Student Work

Student work also provided data reusers with information about the context of class activities as well as insight into student thought processes. For example, Interviewee_025 was interested in getting a full perspective on class activities and noted that student work examples gave ‘...a very clear sense of what kind of problem or task they're working on. So I guess it's more the stuff about making sense of what's going on in the classroom in the video, more than like general background information.’

Interviewees were very interested in student thought processes, both for research as well as teaching aims. Interviewee_039, a researcher, summed up this desire: ‘It would be very nice sometimes to have, in addition to students' work, to have interviews with children that were part of a certain lesson [to] get a better insight into the children('s) perspective on the learning process.’

For teacher-educators, student work was used in two ways. One was to instruct pre-service teachers on the different ways students made sense of problems and thus reflect on the ways teachers could better direct and support students, and second, to encourage student teachers to think about their own thought processes. Student work helped pre-service teachers engage with the problem. ‘What I find very effective is if I can look at classrooms of kids doing problems that my own, the people I'm working with, that we've also been able to engage with the problem’ (Interviewee_033). Likewise, Interviewee_036 discussed the centrality of student work in her teaching,

‘We analyze student work in our numbers and operations course every lesson, but which kind of student work depends on the lesson and in that course the focus is on understanding why a student did what they did and determining what things are mathematically legal and what things are not mathematically legal and how one would justify those and if justifications are sufficient.’

However, for her, the articulation of student thinking also honed the ability of pre-service teachers to develop the skill of analyzing student work,

‘...it's dealing with analyzing student work. The use of student work... is to highlight common errors, common misconceptions. So often incorrect work is shared to highlight important mathematical issues and to get them [pre-service teachers] questioning why we do things the way we do them.’ (Interviewee_036)

Finally, Interviewee_011 stated, ‘what's really important is that you can always hear the kids and if they are referring to something that's in writing, having that artifact.’

Demographic Data about Students and Classrooms

Demographic data about the students and the classrooms, often discussed interchangeably, were important to researchers and educators (including both teacher-educators and their students) who are interpreting the behaviours and teaching practices recorded in the video records of practice. Demographic data, such as whether students are English language learners, had special needs, or came from a particular cultural or socio-economic background, helped viewers to understand facets of the classroom interactions or teaching practices in the video. Interpreting teaching and learning *in context* was more readily facilitated by understanding elements of students’ and teachers’ backgrounds—as well as the school or classroom environments—when viewing the recordings, reading transcripts, or examining student work. For example, one teacher-educator commented,

‘Well, I guess you'd consider demographic, but sometimes it matters to know whether it is a charter school, a school of choice, a school where you know that the kids and their families have committed themselves to being there and you don't have some of the challenges that you have in a more typical diverse school where you have motivational challenges and issues’ (Interviewee_009)

In the original survey, respondents were not asked about their use of demographics data; however, in the write-in responses about additional complementary data used, 9 researchers and 4 teacher-educators indicated that they used student demographic data for their work, as well as 2 respondents who used student demographics for both research and teaching. These 15 reusers voiced the most common written-in response to the question about additional data, though none of the responses included an explanation about how the data were used.

In comparison, classroom and student demographic data were second most discussed type of data after transcripts by interviewees. 14 researchers and 17 teacher-educators used demographic data and an additional 5 (3 researchers and 2 teacher-educators) would use demographic data if they were available with the video records of practice. Researchers, particularly those doing comparative studies, wanted demographic data because it provided richer data for analysis and gave further insight into the context of classroom activities, particularly when comparing across schools and classrooms. For example, Interviewee_017 explained,

‘Well, I mean, when you can get it, like it is something that I always try to know whether I have that information or not I try to, at least, make a crass look at the demographics just to see like what does this conversation represent? Like who's sitting at the table? Kind of thing. So, yes demographics are good to have. And I think again that speaks to understanding the context as much as possible.’

Teacher-educators additionally found demographic data informative when selecting video records of practice to use with their student teachers because they wanted to provide examples of teaching interactions where the student demographics were similar to what the student teachers would experience in their placements.

In a perfect world, if we can find students and/or teachers that maybe the demographics of their school or their student population in the video is relatively consistent with what the students in our class see here locally, that's then again a kind of wishful thinking...That way you kind of you could watch

video tapes or watch clips that are consistent with maybe a placement or teaching assignment that you have locally. (Interviewee_019)

Other teacher-educators prioritized student diversity when selecting videos, although it was often difficult to find videos which demonstrated the desired diversity.

I prefer to have classrooms that show diverse students. One of the things that we just struggle with in general is exposing our teacher candidates to classrooms that are diverse, and so I often try to find videos that do that 'cause we can't always find a real classroom that's diverse, so that's part of it. (Interviewee_010)

We have provided some examples of the types of complementary data teacher-educators and researchers used in conjunction with video records of practice. In the next section we discuss what happens when these are not available.

Desired Complementary Data

Complementary data is not always available in the same or another repository. Nine interviewees (3 researchers and 6 teacher-educators) identified transcripts as the primary type of complementary data they would like to have but were not always available alongside the video records of practice. Several acknowledged the time required to make quality transcriptions of classroom interactions as a reason why they were not available; however, their need for transcripts aligns with the reasons for use outlined above: a clearer record of interactions for interpretation.

The other most desired but not always available data included classroom and student demographic data. In addition to being the interviewees' second most used type of complementary data, they were also the second most discussed data that were not always available with the classroom videos.

'Yeah, I mean, so I think ours is, maybe not unique, but maybe unique to other things where we're looking to have demographics at a minimum and even better would be outcome measures and other stuff like that. And so really looking around and talking to people, there were just almost no repositories of existing data where we could do that. So, that was sort of the first criterion.' (Interviewee_012)

Similarly, curriculum materials provided additional contextual information about the entire arc of classroom interactions and activities longitudinally. These placed video records of practice, which capture moments in time, in context rather than a vacuum; demonstrating how the video is a continuation of or precursor to other activities. Knowing the larger curricular context allowed reusers to situate the videos within a larger classroom context: What students had or had not already been taught; where teachers may have deviated from the curriculum guidelines; and what challenges there might be with the material.

'Having access to, whether it's curriculum materials or it's something about what the plan was that the teacher had, we found that to be pretty essential. We've had some data before where we didn't have access to that...We didn't really know what was going on and if you don't have some sense of what the teacher was intending or what the teacher was working with and what the actual problems were that the teacher had, that she was giving to the students or something, then it becomes very hard...' (Interviewee_004)

When curriculum materials are not available, the video records of practice were often not useful because of missing-but-critical context.

In addition, researchers wanted other types of data that were often not naturally occurring in the classroom, such as teacher journal entries or interviews with teachers where they articulated their intent and understanding of the classroom environment.

‘I particularly find teachers' diaries or teachers explaining what they think were going on. Even if I don't think the same thing is going on, I would like to know what the teachers' perspective on what they think was going on in that interaction or what they were trying to do with the students or, you know, something about the teachers' way of thinking about the interaction would be helpful.’ (Interviewee_018)

Finally, in addition to exploring the types of complementary data that were important for or desired to support reuse—and the various motivations reusers had for needing this data—we examined how these data were used together, in combination, to provide different perspectives and offer a fuller, richer, more complete narrative of the videos. Two themes from the interviews are discussed below: co-occurrence and completeness.

Co-Occurrence

Throughout the interviews, several types of data co-occurred. Co-occurrence signifies relationships between themes as well as the prominence of specific combinations of qualitative codes (Guest, MacQueen, & Namey, 2012). Out of 109 references to transcripts, they co-occurred 71 times with other complementary data types, most commonly: lesson plans (12), class/lesson materials (10), student work (10), and classroom demographic data (9). There was also overlap among many of these co-occurrences, as researchers and teachers discussed the value of having richer data to allow for triangulation among multiple data sources. Interviewee_033 specifically discussed the intersection of data types in making sense of video:

‘What I find very effective is if I can look at classrooms of kids doing problems that my own, the people I'm working with, that we've also been able to engage with the problem. Having the task is really helpful. Having the student work so that we can analyze student work. All these different ways to get at some sort of core activity so they can do the activity themselves...Yeah so transcripts, student work. I think what would be really interesting is lesson plans for the activity. I haven't seen those anywhere but I think that would be really interesting. What did the teacher think about and do for this?...What was the planning that the teacher did before we actually saw the lesson? I think that would be really helpful, don't see that too much.’

Among class materials, lesson plans, student work, transcripts, and classroom demographic data, the first four had the greatest amount of co-occurrence. A comparison of the co-occurrence among these four, as raised in the interviews, is shown in Table 2.

Table 2: Most frequent co-occurrence of mentions of complementary data types

	Class Materials	Lesson Plans	Student Work
Lesson Plans	12	—	—
Student Work	12	12	—
Transcripts	10	12	10

Class materials, lesson plans, student work, and transcripts were types of data that can each be used as data/evidence in studies in their own right. When used together and/or in conjunction with video records of practice, these data provided context of class activities that verified—or shed light upon—the video evidence of teaching and/or learning. For example, Interviewee_036 discussed using the video and student work in tandem to uncover student thought processes, ‘it’s just the classroom videotape, student’s written work, there are a few interviews with students about their thinking that we sometimes use as well.’ The importance of this context was also articulated by Interviewee_021,

‘Primarily things that help us make sense of the video, like our copies of what the students are doing at the time [class materials]; teacher’s lesson plans, so we have a context for what’s going on; pictures of things that happened in the classroom, so things like writing on the board [class materials].’

Completeness

Overall, interviewees viewed the addition of complementary data as adding perspectives to the video records of practice and a completeness to the narrative of the video. For example, Interviewee_001 noted,

‘Some of it is partly redundant. It’s not redundant in the sense that something can substitute for something else, but it’s sort of seeing and interpreting things at different stages in time and with more and more information available. So, it’s a kind of opening up in a much more expansive way.’

Likewise, Interviewee_017 explicitly tied this idea for complete data to data reuse,

‘I mean it’s great because there’s teacher notes from that day, there’s a lesson, there’s like what was happening in the lesson, there’s even the beginnings of a transcript. So, like, that all has been very helpful in trying to at least understand the context, ‘cause, I mean that’s the hard part about using video that you didn’t collect. It’s already hard to understand the context in the place where you are the field researcher.’

In general, data reusers of video records of practice felt there was always something just out of view of the camera lens. As Interviewee_013 summarized,

‘The last piece I would say is just comprehensiveness of their data corpus...But just beyond having the video of the classroom, having the lesson plans, having the seating charts, and...it’s really important we know who is saying what. So to be able to have passed that on to a third-party transcriber for example, and having all of that information for them to go through as they build a transcript that saves us a lot of time.’

Therefore, having a rich array of data to triangulate or present in a teacher education class enriched the research or teaching activity: As one teacher-educator commented, ‘I think that they add value to my experience as a teacher-educator because they give me a lot of materials to choose from and think about’ (Interviewee_027).

Discussion

The results of this study have several implications for the data collection practices of the wider data repository community, not just those who work with video records of practice. This study found that, while those using video records of practice for research or teaching identified the same types of complementary data, they often met different needs. There is an overarching desire, however, of both researchers and teacher-educators to have richer, more complete data that provide additional context for video records of practice. Furthermore, our survey respondents and interviewees provided a very rich list of the types of complementary data they needed, as well as whether they were able to find those data. These data types go beyond what many repositories typically expect to collect when they bring in new data sets or may think to collect when accessioning video records of practice.

These different needs, as well as different motivations for using the same types of complementary data, provide guidance for future repository data collecting, referral to related data, and metadata application. When considering the accession of new data collections, especially those intended to be used by a broad community of reusers like the researchers and teacher-educators described here, repositories need to identify what complementary data might provide additional context for the target data. Then, they must determine what comprises a more complete data universe—what exists and might be collected for inclusion in the dataset—and what is available through other sources and might be linked within the metadata description.

Conclusion

In this study, reusers of educational video records of practice articulated how critical complementary data were for their motivations and intentions for data reuse. The exploration of these rationales for wanting complementary data and how data reusers intended to use these data add to our understanding of the intricacies of data reuse. In particular, many of the complementary data types desired when using video records of practice were used to give meaning to events and offer insight into the decisions behind activities seen on the videos. While metadata is an important and valuable means of describing data for reuse, meaning is often constructed through comparison, verification, or elucidation in reference to other data. Repositories that make data available for reuse should be aware of how complementary data are utilized to meet the needs of their reusers.

Acknowledgements

This project was made possible in part by the Institute of Museum and Library Services grant # LG-06-14-0122-14.

References

- Andersson, E., & Sørvik, G. O. (2013). Reality lost? Re-use of qualitative data in classroom video studies. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 14(3). Retrieved from <http://www.qualitative-research.net/index.php/fqs/article/view/1941>
- Dempsey, L., Heery, R., Hamilton, M., Hiom, D., Knight, J., Koch, T., ... Powell, A. (1997). *Specification for resource description methods. Part 1. A review of metadata: A survey of current resource description formats*. (No. RE 1004 (RE)). Retrieved from UKOLN, University of Bath website: http://www.ukoln.ac.uk/metadata/desire/overview/rev_ti.htm
- Faniel, I. M., Frank, R. D., & Yakel, E. (2019). Context from the data reuser's point of view. *Journal of Documentation*, 75(6), 1274–1297. <https://doi.org/10.1108/JD-08-2018-0133>
- Gilliland, A. J. (2016). Setting the Stage. In M. Baca (Ed.), *Introduction to Metadata* (3rd ed.). Retrieved from <http://www.getty.edu/publications/intrometadata/setting-the-stage/>
- Guest, G., MacQueen, K. M., & Namey, E. E. (2012). *Applied thematic analysis*. Thousand Oaks, C.A.: SAGE Publications.
- Hartwood, M., Procter, R., Taylor, P., Blot, L., Anderson, S., Rouncefield, M., & Slack, R. (2012). Problems of data mobility and reuse in the provision of computer-based training for screening mammography. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems Pages*, 909–918. <https://doi.org/10.1145/2207676.2208533>
- Hatch, T., & Grossman, P. (2009). Learning to look beyond the boundaries of representation: Using technology to examine teaching (overview for a digital exhibition: learning from the practice of teaching). *Journal of Teacher Education*, 60(1), 70–85. <https://doi.org/10.1177/0022487108328533>
- Hedstrom, M., & Lee, C. A. (2002). Significant properties of digital objects: Definitions, applications, implications. *Proceedings of the DLM-Forum: @ccess and Preservation of Electronic Information: Best Practices and Solutions*, 218–223. Barcelona, Spain.
- Hedstrom, M., Lee, C., Olson, J., & Lampe, C. (2006). 'The old version flickers more': Digital preservation from the user's perspective. *The American Archivist*, 69(1), 159–187. <https://doi.org/10.17723/aarc.69.1.1765364485n41800>
- Lee, C. A. (Cal). (2011). A framework for contextual information in digital collections. *Journal of Documentation*, 67(1), 95–143. <https://doi.org/10.1108/00220411111105470>
- Wallis, J. C., Rolando, E., & Borgman, C. L. (2013). If we share data, will anyone use them? Data sharing and reuse in the long tail of science and technology. *PLoS ONE*, 8(7), e67332. <https://doi.org/10.1371/journal.pone.0067332>