

4-12-2021

YouTube for Transcribing and Google Drive for Collaborative Coding: Cost-Effective Tools for Collecting and Analyzing Interview Data

Tim Hopper

University of Victoria, thopper@uvic.ca

Hong Fu

University of Victoria, fuhong@uvic.ca

Kathy Sanford

University of Victoria, ksanford@uvic.ca

Thiago Hinkel

University of Victoria, thiagohinkel@uvic.ca

Follow this and additional works at: <https://nsuworks.nova.edu/tqr>



Part of the [Educational Assessment, Evaluation, and Research Commons](#), and the [Educational Technology Commons](#)

Recommended APA Citation

Hopper, T., Fu, H., Sanford, K., & Hinkel, T. (2021). YouTube for Transcribing and Google Drive for Collaborative Coding: Cost-Effective Tools for Collecting and Analyzing Interview Data. *The Qualitative Report*, 26(3), 861-873. <https://doi.org/10.46743/2160-3715/2021.4639>

This How To Article is brought to you for free and open access by the The Qualitative Report at NSUWorks. It has been accepted for inclusion in The Qualitative Report by an authorized administrator of NSUWorks. For more information, please contact nsuworks@nova.edu.

A promotional banner for the Qualitative Research Graduate Certificate at Nova Southeastern University. The banner has a dark blue background on the left with the NSU logo and text: "Qualitative Research Graduate Certificate", "Indulge in Culture", "Exclusively Online • 18 Credits", and a "LEARN MORE" button. On the right, a group of six diverse people are sitting on a stone ledge in front of a building with "NOVA SOUTHEASTERN" visible on the wall.

Qualitative Research Graduate Certificate
Indulge in Culture
Exclusively Online • 18 Credits
LEARN MORE

NSU
NOVA SOUTHEASTERN
UNIVERSITY

NOVA SOUTHEASTERN

YouTube for Transcribing and Google Drive for Collaborative Coding: Cost-Effective Tools for Collecting and Analyzing Interview Data

Abstract

Cloud-based tools are increasingly used in research processes. In this paper, we illustrate the practice of one research team making use of multiple cloud-based applications in preparing, analyzing, and sharing research data, as well as in collaborative writing and display of results. Important research ethics considerations are also explored as a foundation for this practice. We believe that our detailed description of the steps involved can be of help to researchers, particularly novice researchers who may lack research funds to have qualitative interviews transcribed. This mashed-up use of free cloud-based software makes data preparation from qualitative interviews cost-effective, more efficient, thorough, and collaborative.

Keywords

data preparation, cloud-based tools, collaborative writing, protection of privacy

Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial-Share Alike 4.0 License](https://creativecommons.org/licenses/by-nc-sa/4.0/).

YouTube for Transcribing and Google Drive for Collaborative Coding: Cost-Effective Tools for Collecting and Analyzing Interview Data

Tim Hopper, Hong Fu, Kathy Sanford, and Thiago Alonso Hinkel
University of Victoria, British Columbia, Canada

Cloud-based tools are increasingly used in research processes. In this paper, we illustrate the practice of one research team making use of multiple cloud-based applications in preparing, analyzing, and sharing research data, as well as in collaborative writing and display of results. Important research ethics considerations are also explored as a foundation for this practice. We believe that our detailed description of the steps involved can be of help to researchers, particularly novice researchers who may lack research funds to have qualitative interviews transcribed. This mashed-up use of free cloud-based software makes data preparation from qualitative interviews cost-effective, more efficient, thorough, and collaborative.

Keywords: data preparation, cloud-based tools, collaborative writing, protection of privacy

Introduction and Context

Technology has been increasingly used for research purposes such as collecting, storing, and analyzing data, as well as collaboratively writing research findings. For example, Given and Willson (2018) described ways to use and create technology tools for data preparation in decision-making and meta-level processes. This use of technology tools has been enhanced and made more widely available by the expansion of cloud-based applications. A cloud application is a software program that relies on remote servers for processing logic that is accessed through a web browser with a continual internet connection. Cloud application servers are located in remote data centers operated by third-party cloud services infrastructure providers and typically encompass tasks such as email, file storage and sharing, word processing, and other data entry and inventory managing tasks. This type of cloud computing supplies on-demand availability to an array of software tools free or for low cost and with high levels of stability when continuous internet connection is available. Data stored on cloud services is instantly available to authorized users in any location globally. Due to their massive scale, cloud providers can hire world-class security experts and implement infrastructure security measures that typically only large enterprises can obtain.

Taking advantage of the security, processing power, and access, researchers such as Stockleben et al. (2017) are creating virtual spaces for collaborative writing where cloud-based applications such as Google Docs appear to be the de facto standard. In such situations, research teams have developed their own style of commenting and marking collaborative documents. In this way, a shared working space using free cloud-based tools can be mashed together to create sophisticated systems of coding and analyzing where “no one acts as administrator or gatekeeper” (p. 585). It is evident that cloud-based systems and apps are becoming more commonplace for researchers in their daily activities.

Building on this emerging use of cloud-based technologies, in this paper we outline a cost-effective process for collecting, coding, sharing, and analyzing large quantities of interview text. Typically, in Canada, a one-hour interview with a participant will take a skilled transcriber four hours or more to transcribe at a cost of approximately \$120 or above. Once a transcript has been completed, researchers will share it with the participant to check for accuracy (commonly referred to as member checking) before attempting to code and analyze the transcript as data, often comparing multiple transcribed interviews; this process is costly and time consuming. For researchers without funding, the transcription process can be a daunting endeavor.

For the past several years, our research team has been investigating how digital technology can be used to improve students' learning experience and teachers' assessment practices (Hopper et al., 2016, 2018; Sanford et al., 2013; Walker et al., 2017). Therefore, digital and cloud-based apps have been integral to both the content and the process of our research work. While looking into participants' use of digital technology in educational activities, we have found that our research practices can also benefit from these digital affordances. Steps in qualitative analysis processes can all be carried out with higher efficiency, for less cost, and yielding quicker and better results thanks to cloud-based digital apps. Particular highlights of this cloud-based research process are: (1) the ability to video record an interview (capturing the contextual elements as well as the participants insights), (2) using a private YouTube channel to generate the text of an interview using the closed captioning feature in YouTube for no cost, then shape this text into a transcript in the same time as the interview, (3) sharing transcripts with participants as a Google document for edits and comments, (4) coding of the finalized transcript using color coding to develop categories collaboratively with the research team, and (5) transferring the coded transcript to NVivo for final coding based on emerging themes from the collaborative process.

One issue that has been raised when we shared this research analysis process is the ethics and security of sharing participants' personal data through cloud-based systems. Concerns are raised, for example, over who has access and government legislation such as the Patriot Act in the US that, in theory, means the US government can access personal information if they believe there is a terrorist threat. However, this access is limited, with access only allowed to targeted individuals¹. In the next section we will address some of these security and ethical concerns. After that we outline five steps for recording and generating transcript-based data via: (1) collecting interview data, (2) transcribing the interview, (3) sharing with participant to check accuracy and intent, (4) collaboratively coding, and then (5) developing analysis leading to themes for publication.

Security and Ethical Considerations in the Proposed Cloud-Based Research Process

Exploring research ethics in virtual spaces, including cloud-based spaces, has been a recent development, with only some broad discussions but scarce empirical research. When researchers increasingly use these spaces for research purposes, a key concern that emerges is security and privacy. According to Asher et al. (2013), security issues in virtual spaces present a serious conflict for scholars, as maintaining data security and privacy is central for researchers to satisfy both their own professional codes of ethics and the confidentiality standards established by institutional review boards (IRBs), which are unlikely to review cloud-based systems third-party end user agreements for privacy risks. Furthermore, many of the security concerns surrounding cloud computing systems are not unique to commercial cloud providers (e.g., Amazon, Microsoft) but may be characteristic of the way cloud networks are currently

¹ https://en.wikipedia.org/wiki/Patriot_Act

constructed. Therefore, an important practical challenge within the cloud computing environment is to develop the ability to note any risks to participants and to note what service providers indicate in their service agreements. Currently there is a significant gap in internet research ethics on the benefits and cautions of using cloud-based apps; however, with the high cost-benefits of using these systems, just saying “no” to using such powerful tools is not a solution. As noted by Warrell and Jacobsen (2014), educational and social science researchers in Canada should “continue to seek better and more responsive guidelines that describe how to deal with ethical issues arising in various online research situations, contexts, and conditions and to educate the community by publishing and sharing their practices, findings, and solutions” (p. 35). In what follows, we explore these security and ethical issues in relation to the cloud-based research processes before sharing our practices using these cloud-based technologies.

YouTube and Google Password Protection Security

Password systems are used on an array of computer-based, on-line and cloud-based systems to protect access to personal information and to identify users for services they have selected and sometimes paid for. A hacker is a person who tries to circumvent password protections to access individuals’ accounts, personal information, and services. The YouTube and Google Drive are protected by Google’s two-factor authentication password where user identification and password are associated with a person’s devices. In contrast, password systems for most internet and university-based systems are just a username and an 8-character or more password. From a password point of view, YouTube and Google Drive are more difficult for a hacker to access.

The YouTube system is a powerful tool for storing video data. Video data is stored in Google Data servers known as data centers distributed in 19 locations in the US, 12 in Europe, one in Russia, one in South America, and three in Asia. These data centers are server farms that host thousands of servers with hundreds of thousands of hard drives. They have a level of security that no other company or nation can possibly match. The locations of those server farms are top secret to prevent hacker attacks or terrorist attacks (Google Workspace, 2014). Similarly, other cloud-based systems, such as Microsoft business services that host the email exchange Microsoft Outlook (formerly Hotmail), have multiple data centers around the globe with many layers of security to protect customer information (Microsoft, n.d.). From a password point of view, it is harder to break into or hack a person’s YouTube and Google account than it is to hack into a person’s email.

YouTube as Transcriber of Interviews vs. Emailing Transcripts to Participants

Ethical issues of using YouTube to generate a transcript from a research interview can be divided into three key areas: (1) where personal information is stored; (2) how transcripts are shared with participants after being generated; and (3) what is done with data stored on cloud-based servers. The first issue is concerned with the Freedom of Information and Protection of Privacy Act (FIPPA). In British Columbia for example, FIPPA sets out the access and privacy rights of individuals as they relate to the public sector, such as university, schools, and public services, in regard to personal information that is stored in cloud-based servers. This means that individuals have control over how their personal information is stored and accessed. In the case of research with human subjects, FIPPA indicates that participants have the option, once they have volunteered to be a participant in a study, as to how their information is accessed, shared, and disseminated. In the case of using YouTube to generate a transcript, participants ethically need to have the option to have their interviews generated by a transcriber

or by using the YouTube closed-caption feature. The key requirement is that participants have a choice.

In regard to the second issue, once a transcript has been generated, it is commonly shared with the participant through email. This is a convenient method. However, in the same way that YouTube is stored in Google data servers around the world, most email communications, depending on the email service, are stored in similar data servers. These systems use an IMAP (Internet Message Access Protocol; “Internet message access protocol,” 2021), which downloads emails and leaves them on the server, unless the user deletes them deliberately. This means one can download the same emails to several PCs or read them online with a different smartphone, tablet, PC, etc. However, like YouTube, the emails and their attachments are held on servers. If a person is using Gmail, then the email is held on the same data servers as YouTube. Most researchers email participants their transcripts for them to read and check, so this means the transcript (generated by a transcriber or by YouTube) sent via email is being stored on data servers located all over the world. As can be seen by many well published leaks of politicians and celebrity emails, these email accounts seem to be vulnerable to a motivated hacker.

The last issue relates to what is done with personal information stored on data servers. The cloud-based providers can scan data stored in cloud-based servers unless they are paid for by the user. For example, Google artificial intelligence (AI) scans data stored on online servers in email, YouTube, and Google Drive. Before 2017, Google scanned and monetized email accounts, such as personalizing services to a customer (Schofield, 2018). However, the company announced in 2017 that they would no longer scan emails to tailor adverts (Simon-Lewis, 2017). It is complicated to assess what is being done with personal information held by different service providers, but it is difficult to avoid personal information being accessed even if not associated with a person. Through an email message, information is associated with the person by virtue of their email address; information on YouTube can be visual but not associated with a person unless specifically labeled.

Summary of Security

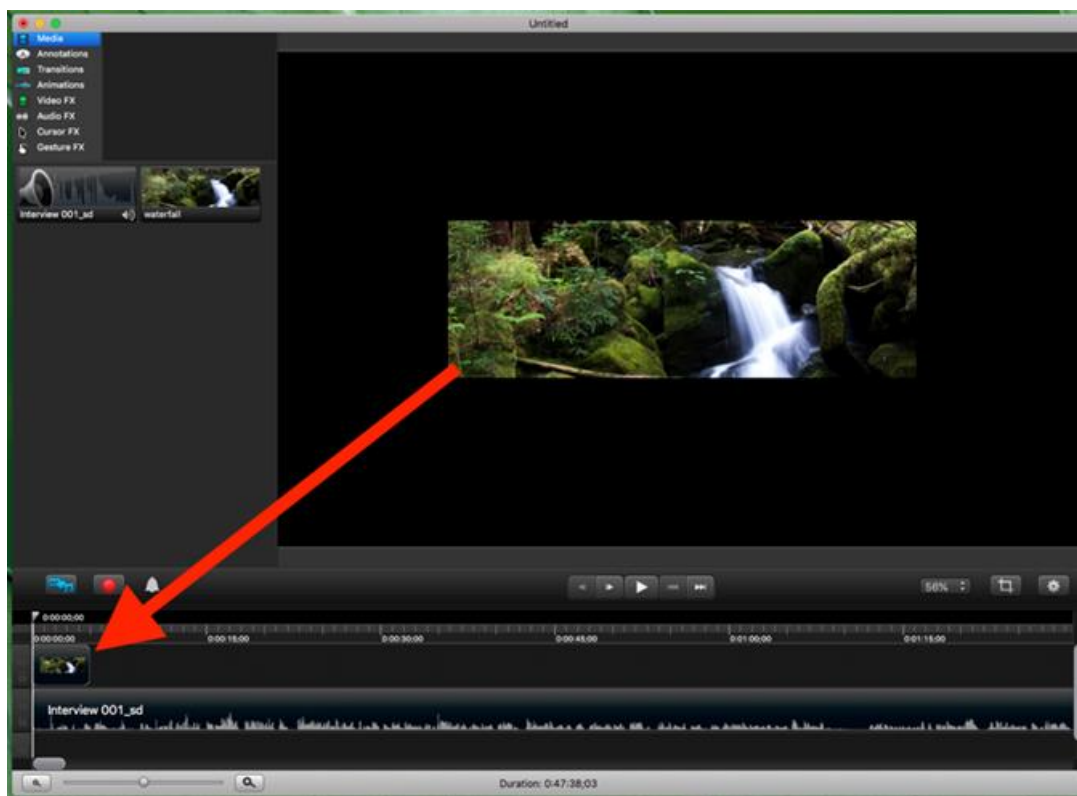
The key message we can assess from this section on security and ethical issues related to using cloud-based applications is caution. It is important to check what is being done with the data by the service provider and to be explicit with research participants so they can make informed decisions. However, it is important to recognize that using YouTube and Google Drive is not less secure than sending a transcript via email to a participant. Depending on the email system the participant is using, this transcript information could be scanned and used for services that make the email system free but not totally private. Ideally, a researcher would conduct an interview, transcribe it without using a cloud-based system, and then send it by mail, more digitally secure but clearly this could be physically intercepted before reaching its destination; it would also be costly and very time-consuming. However, in such a process, a transcript would not be scanned without the researcher’s knowledge. We believe using cloud-based systems allows the researcher, especially those such as graduate students with limited funds and time, to access critical insights from their transcribed interview data in a timely and cost-effective manner. The sharing of transcripts quickly with participants also leads to more engaged member checking where, in our experience, the participant feels acknowledged and more connected with the study as they receive the transcript within a day or so of the interview, which can lead to additional insights by participants in any follow-up debriefing with them.

Preliminary Coding Process Through YouTube and Google Docs

Digital technology plays a vital role in data collection of our research process. For example, we conduct semi-structured interviews with our participants, which are either recorded as audio or video depending on the participants' preferences and research intents. As described by qualitative researchers, a more conversational style of interviewing (Creswell & Poth, 2017; Sparkes & Smith, 2014) can be developed and directed by a few leading questions. This style of interviewing can provide sufficient room for participants to expand on points of interest to help construct the meaning making process. We often end up with interviews of various lengths, from approximately 25 minutes to one hour in some cases. Obviously, transcribing long interviews can be time intensive, which can discourage sustained conversation between researchers and participants.

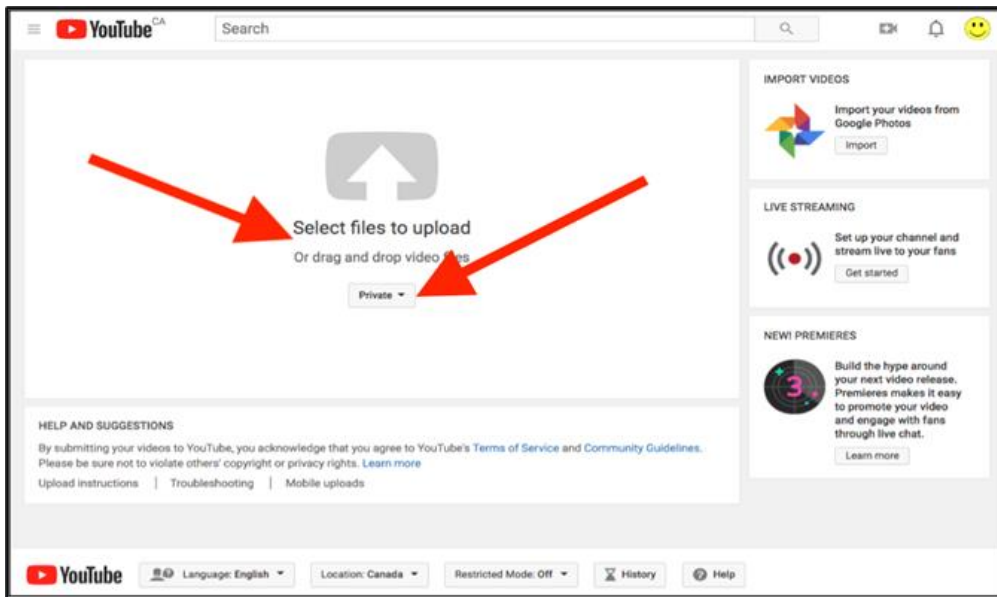
The YouTube closed captioning feature translates the audio text into on-screen text. This feature can be adapted to provide an efficient way to generate large quantities of text data from interviews. Specifically, with informed consent from participants to upload interviews as private videos, we use the YouTube auto closed-caption function to generate the words of a transcripts, which are then "cleaned up," identifying who spoke the words and adding clarifying punctuation. This job is done by a member of the researcher team or a transcriber within a shorter time than typical transcribing, usually just over the time it takes to complete the interview. As YouTube does not allow mp3 audio files to be uploaded, in case of audio interviews the video creation applications in Windows or Apple operating systems can easily transform an audio interview in mp3 to an mp4 video by just adding an image as "cover," shown in Figure 1. Once uploaded to our research account as private videos, it takes approximately 30-45 minutes for the auto-generated caption to appear (see Figure 2), depending on the length of the video file.

Figure 1



Note. Audio iPad interview converted to mp4 by adding image using Camtasia program.

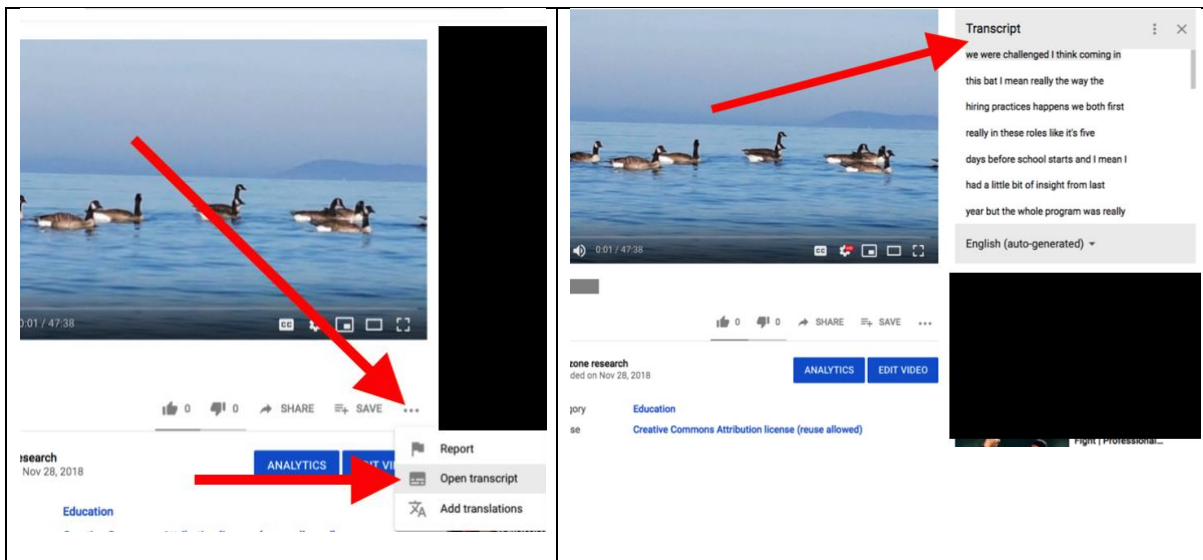
Figure 2



Note. Uploading to YouTube private account by dragging mp4 file into window.

The next steps, shown in Figure 3, are to open transcript options (see "...") and extract the transcript from YouTube (see "open transcript") and then copy the text in order to paste it into a word-processor program like Microsoft Word or Google Doc. This transcript text can be toggled to show "time-stamp" or "no time-stamp" if you just want the text. Note the text in the transcript is hyperlinked to when spoken in the video. This is very useful to relocate important phrases in an interview and can be aided by keeping a time-stamped copy of the interview.

Figure 3

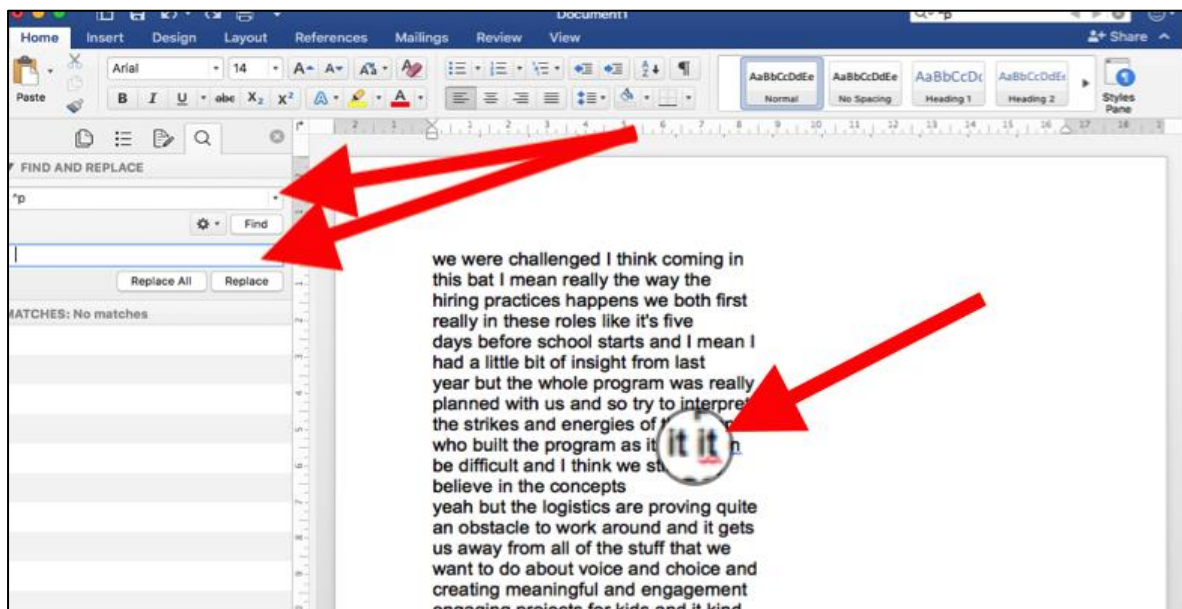


Note. Extracting text and copy the generating text.

In Figure 4 the transcript copied from YouTube is pasted into Microsoft Word to be converted from a column of text with line-breaks into a transcript of the interview words to be edited into paragraphs, separated by who spoke the words (see Figure 4). At this stage, the transcript can be converted from a column format (with new-line breaks added by YouTube)

to a stream of text that can be edited into a transcript format. This conversion is done in Microsoft Word by doing a search and replace feature of changing paragraph mark shown as “^p” code to be replaced by an empty space. At this time obvious edits such as double words shown by red underline can be fixed (see “it it” example in Figure 4).

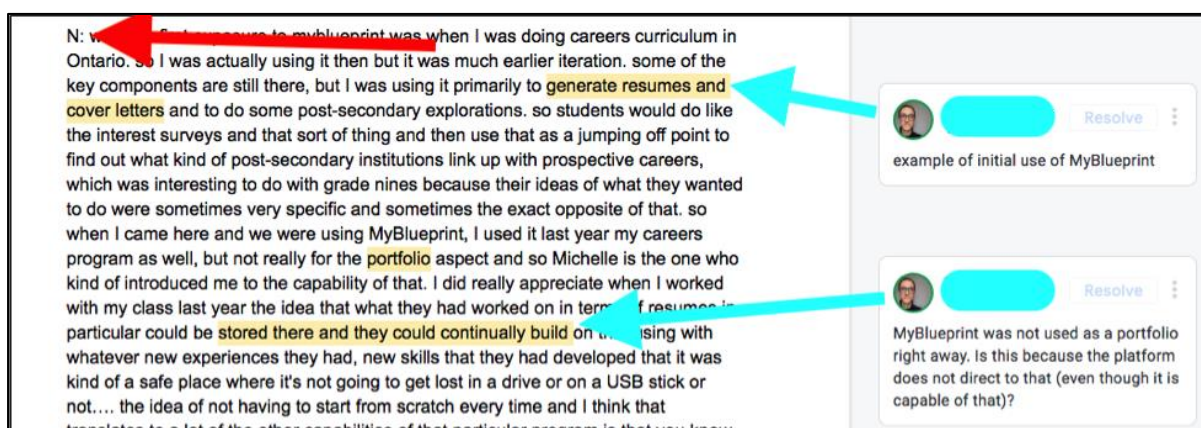
Figure 4



Note. Pasting of transcripts from YouTube into the word processor.

As shown in Figure 5, the completed transcript, edited to indicate who was speaking (see “N:” in Figure 5) can then be shared via Google Docs to the research team, and links to Google Docs are sent back to respective participants for member checking with possible comments and clarifications from the researcher. The whole process of transcribing, editing, member checking, and commenting happens within a short time span, which works well for ongoing conversations between researchers and participants.

Figure 5



Note. Edited transcript sent to participant and shared with research team with comments.

Google Docs also enables the research team to collectively do initial coding by identifying salient narratives, key concepts, and recurring ideas. One researcher can start off

the preliminary color coding by reading through the transcripts and identifying narratives/concepts that are shared or distinct (See Figure 6). A useful Google add-on tool we have discovered that helps with this color-coding process in Google Docs is called “Highlight Tool,” created by a high school student in 2015 and updated based on user feedback as he became a college student.² The add-on extracts coded text at end of the document into a separate document grouped by color or sequenced by color. This is a very effective way to group coding in clusters that can then put into a separate document and coded in connection to a category. The research team members can then all access the transcripts to add in their comments and offer input on the same transcripts, using the feature of Google Docs for collaborative writing and editing. An index, as shown in Figure 6, is then created that can be built into the Highlight Tool that corresponds to the text coded with the label created by the research team. For example, in Figure 6 the green for “ongoing/self-assessment” label corresponds to text in the transcript that is coded green, which connects to this idea or to the coded pink section for “share their learning” quotes.

Figure 6

The screenshot shows a Google Docs interface for a document titled "Interview". The toolbar includes a red arrow pointing to the "Text colour highlight feature" (represented by the letter 'A' with a color swatch). Below the toolbar, a legend lists various categories with corresponding color swatches: Student ownership (red), Purpose (orange), Ongoing/self assessment (green), Communicating with parents (and other stakeholders) (blue), Repository (Archiving learning) (purple), Share their learning (pink), Demonstrate learning (dark blue), Curriculum (red), Reporting (yellow), Tech learning (orange), No "guidelines" is good (brown), Challenge (continuous / school wide use) (cyan), Collegial support (green), and Planning activities ahead (purple). A dashed box highlights a transcript snippet with color-coded text: "P: yeah well cause it's metacognition. You've got to think about what you're what you're learning what you're thinking about. I want them to be aware like why is this piece of work your best or why do you like it or what do you want to improve like, that to me is really important because I'm not the only one who's going like I want the kids to have an intrinsic motivation I want them to want to do their personal best and by reflecting on what they've done and writing a comment what did you like about this what was hard for you right what did you find challenging what did you find interesting, asking them think about it. I find the quality of work that they produce is better and knowing that they have to upload this work onto fresh grade for their parents to see, I find it's my hope that they will then improve upon it and it's work either parents will see something and they'll be like yeah can so-and-so redo this or parents will or they'll know their parents are watching so (yeah they have the audience) yeah the audience built-in yeah". A comment box on the right says "also fit for 'student ownership'". Below the legend, the transcript continues with: "H: yeah no I'm talking with [redacted] about her practice of using fresh grade using digital technology for communicating student learning and for assessment okay so how long have you been actually yet some background like how long have you been teaching elementary? P: twenty-one years H: oh that's quite a long time and and you told me just now that you've been using fresh grade for one year?"

Note. Color coding to share and review which allows core-coding by team with comments.

² Website on add-on tool: https://jsonchin.github.io/highlight_tool/

The color-coded labels are then entered into a matrix of components in the research, shown in Figure 7. This matrix of constant comparison is referred to as componential analysis by ethnographic researchers (Spradley, 1980). In this case, the components refer on the x-axis to participants, and the y-axis represents domains and emerging categories or themes from the codes. Representative anecdotes and quotes from participants are then imported into the matrix to illustrate these domains/themes in greater detail. The matrix can be shared with the research team in a Google Docs format, thus offering further tools for collaboratively comparing and contrasting data generated from the interviews.

Figure 7

Emerging domains/themes		Teachers componential analysis				
Domain	category/agent	Names of participants				
Demographic	Teaching time. Experience.		new teacher, cross-curricular	French immersion	grade 6 (math lesson observed)	
Why develop a digital portfolio?	Student ownership / centered	I think they're gonna get a lot more out of it and feel like this isn't just a hoop to jump through, that there's value in themselves and their own lifelong learning journey it's about metacognition and being able to have like you said artifacts to reflect on easily accessible and maintain an awareness of that growth. the online platform really also allows for diversification in a way that sometimes traditional assessments don't the social skills and an awareness of themselves and their own passions and talents are just as important as	and I love that because they learn they get to reflect and they're learning so much more. They get so much more out of it when they do that and there's that conversation with everyone. I find we were able to celebrate so much very exciting to see them go from like I want you to post this project you can use it differently ...	I'm hoping one day that I can get there with fresh grade is here are the competencies that the children need to attain and they actually can pick the processes of how to get it so every kids portfolio looks different you know	whatever they make and choose if they want to make a video or do the work and take a picture or whatever works for them, and then so after they upload the video then they go in and they do they assess themselves and say whether they're meeting expectations and then why and give evidence, and then say something they're proud of, and a goal for next time and I'm getting better at being more flexible with	I guess that participatory element the onus is on the student to be doing most of the work if the teacher is doing most of the work it's not the students portfolio, it's the teacher's portfolio, so there's that piece and yes an overarching understanding that it is about growth not about an end product. It's not a summative experience, it's a growth process. if we could differentiate those types and then if we could leave it to the students to make the decision about what gets kept. The trouble is our young children able to consent meaningfully.

Note. Componential Matrix Analysis of emerging domains and themes.

NVivo Coding Process and Concept Mapping

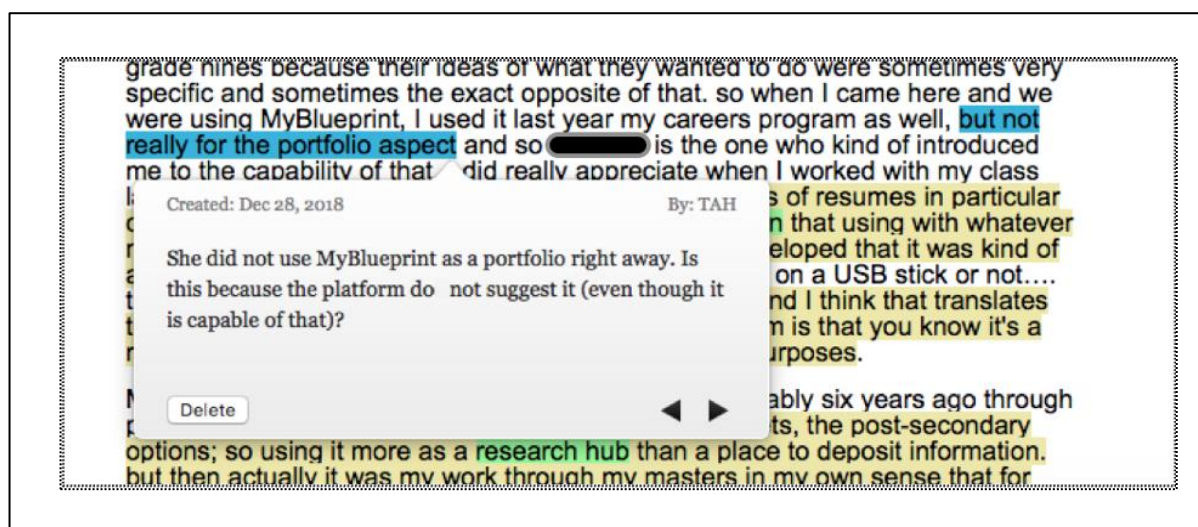
Once the preliminary color-coding process is completed on Google Docs, a team member can then export the transcript and the emerging codes into specialized qualitative data analysis program such as NVivo³. During this step, each interview transcript exported becomes a source. Similarly, each color code representing different narratives becomes a single node or tree node in NVivo. By working with this software, researchers can further explore the emerging categories initially detected and make cross-participant and across-interview connections. Additionally, comments that were added during the process may also be transferred (see Figure 8). These memos serve both as a repository for insights and as a communication channel between team members working with the data in different moments. However, the most valuable characteristic offered by NVivo is its ability to permit users to see the relationship between participant narratives and emerging categories.

In addition, to further identify the details in the emerging themes, researchers can also use NVivo to generate layers of nodes (called trees node) to map the richness of the data.

³ NVivo a purpose-built for qualitative and mixed-methods research <https://www.qsrinternational.com/nvivo/what-is-nvivo>

During this process, categories and meaning emerge from multiple codes across participants to form themes that can also interact with each other. These coding labels (nodes and trees in NVivo) can collect text that links back to the transcript, allowing descriptive definitions to be developed for the emerging category or theme. By spotting similar themes as they work with the data, researchers can tie them under a common category, which can be used to map concepts that relate to those present in the literature. Furthermore, the software has a feature that allows users to obtain statistics of their coding process. By counting occurrences, for instance, it helps identify the relevance of emerging nodes and the respective categories to which they have been assigned.

Figure 8

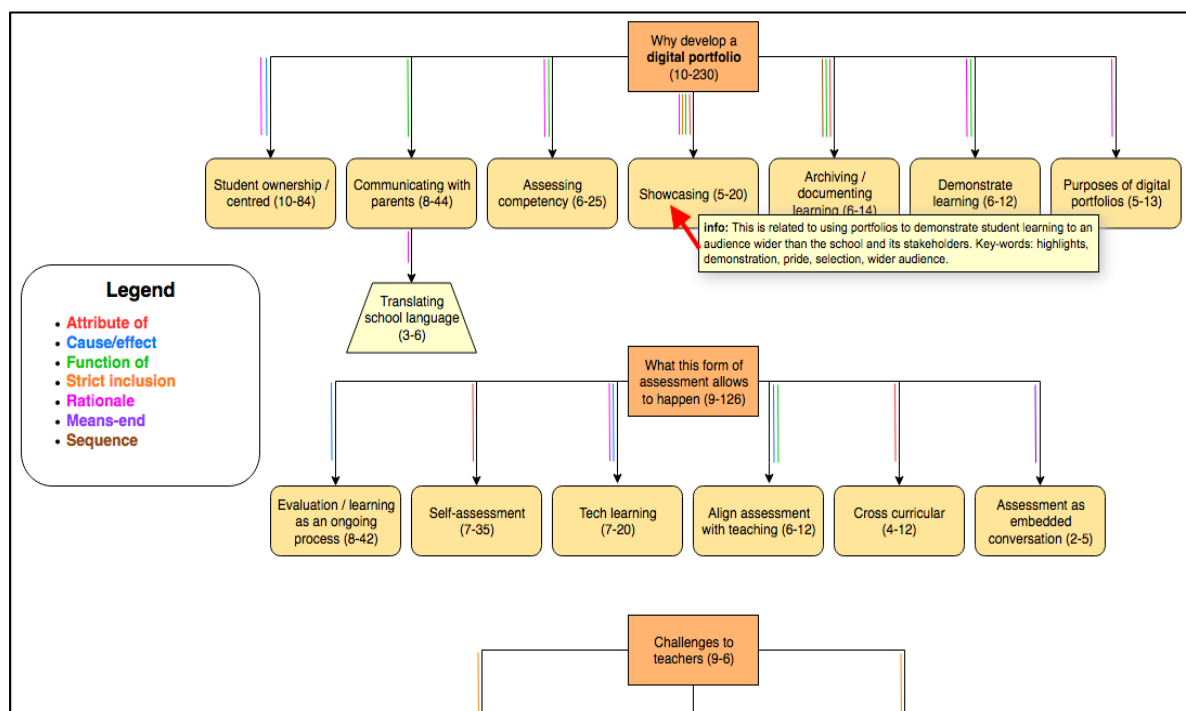


Note. Coded transcript imported into NVivo for cross case coding - Comments from Google Docs transfer in with doc.

We have also discovered that within the Google Drive suite of add-ons, there is a free application called “draw.io”⁴ by JGraph Ltd. that can be used to concept ideas from the research data. This application enables users to easily create diagrams that can hyperlink back to text data in Google Drive. Similar to features in NVivo, this application allows researchers to visualize the emerging categories/themes and the codes under each one of them. In addition, this application permits for the inclusion of descriptive text of relations between elements that can be copied and pasted from NVivo. In Figure 8, the categories that have emerged are grouped into themes such as “Why develop a digital portfolio?” and if you click on a category like “Showcasing” you see the definition of this category. In the brackets for each label are the number of participants who made comments for this category, coded with this label and number of quotes coded. For the purposes of this paper, we aimed at identifying the semantic relationship between codes and their respective categories as well using colors to indicate types of semantic relationship such as “cause/effect,” “sequence,” or “function of,” etc. (Spradley, 1980).

⁴ <https://drawio-app.com/>

Figure 9



Note. Concept Map with emerging themes.

All these on-line tools are free. The use of NVivo adds another level of links to search and index to different ideas spoken by participants. However, the core of the coding is done using cloud-based, freely accessible applications that offer secure ways of developing collaborative coding and analysis. Other software that achieves this transcribing of interviews from digital recordings is available for a monthly cost. For example, NVivo now has a transcription function, Otter.ai allows live transcribing of Zoom meetings⁵, and Dragon Anywhere allows transcribing of mobile recordings⁶. However, unlike YouTube that enables researchers to generate transcripts for free, there is cost attached to using each of these three examples.

Conclusion

In this paper we have highlighted key steps we have discovered when using free cloud-based applications to transcribe, code, share, and develop our analysis of qualitative data. We have noted there is the need to be cautious in how the researcher uses these tools, to check whether privacy needs are being met by third-party end user agreements and to be clear about any risks for any research participants in sharing their personal information, so they can make an informed decision about whether to be involved in the study and how they want the data they generate to be used.

However, in our experience the cloud-based technologies can enable fast and efficient data collection, preparation and sharing in the research process. Access to the cloud-based technologies allows low cost, quickly prepared data collection and analysis that can promote collective research and writing processes within the research team. Access to innovative software applications allows multiple ways for data analysis and result display. We have found

⁵ Live transcribing of Zoom meeting <https://otter.ai/zoom>

⁶ Transcribing of mobile recording in Dragon anywhere <https://www.nuance.com/dragon/dragon-anywhere.html>

these tools very useful by reducing the time delay in accessing interview data. One additional feature of the process we have described, where participants are happy to have their interviews video recorded, is the capacity to see what they are referring to when being interviewed as well as to capture their facial expressions during the interview process. The other feature of a YouTube transcript is it allows word text to be hyperlinked back to the interview. This means text that has gone through a coding process described in this paper can then be searched in the on-line index and linked back to the place in video where the participant stated the coded comments. This can be very useful in adding to or informing the meaning making interpretations in the analysis process. This capacity to link back to original transcripts is very powerful when returning to data after coding and interpreting multiple data sources. In summation, we agree with the review by Stockleben et al. (2017) about cloud-based applications use in inquiry projects, indicating the great potential of mashed-up environments of state-of-the-art web services to be applied to qualitative data processing. These freely available applications can enhance the collaborative capacity and accessibility of doing this type of research, and we see this approach as enabling timely, verifiable, and meaningful data analysis for qualitative research projects for both researchers and participants.

References

- Asher, A., Deards, K., Esteva, M., Halbert, M., Jahnke, L., Jordan, C., Keralis, S. D. C., Kulasekaran, S., Moen, M. E., Startk, S., Urban, T., & Walling, D. (2013). *Research data management: Principles, practices, and prospects*. Council on Library and Information Resources.
- Creswell, J. W., & Poth, C. N. (2017). *Qualitative inquiry and research design: Choosing among five approaches*. Sage Publications.
- Hopper, T., Fu, H., Sanford, K., & Monk, D. (2016). Electronic-portfolio development in three professional programs: Conceptual framework and summary of initial findings. *Journal of Technologies and Human Usability*, 12(2), 13–35.
- Hopper, T., Fu, H., Sanford, K., & Monk, D. (2018). What is a digital electronic portfolio in teacher education? A case study of instructors' and students' enabling insights on the electronic portfolio process. *Canadian Journal of Learning and Technology*, 44(2). <https://www.learntechlib.org/p/184788/>
- Given, L. M., & Willson, R. (2018). Information technology and the humanities scholar: Documenting digital research practices. *Journal of the Association for Information Science and Technology*, 69(6), 807-819.
- Google Workspace [Google Workspace]. (2014). *Inside a Google data center* [Video]. YouTube. <https://youtu.be/XZmGGAbHqa0>
- Microsoft. (n.d.). *Where your data is located*. Microsoft Trust Center. <https://www.microsoft.com/en-us/trustcenter/privacy/data-management/data-location>
- Simon-Lewis, A. (June 27, 2017). Google will no longer read your emails to personalize adverts. *Wired*. <https://www.wired.co.uk/article/google-reading-personal-emails-privacy>
- Sanford, K., Hopper, T. F., & Fisher, P. (2013). Changing paradigms: Embracing contemporary learning theories through ePortfolio. In K. Sanford & T. Strong-Wilson (Eds.), *The emperor's new clothes? Issues and alternatives in uses of the portfolio in teacher education programs* (pp. 72–90). Peter Lang.
- Sparkes, A. C., & Smith, B. R. (2014). *Qualitative research methods in sport, exercise and health: From process to product*. Taylor and Francis Group. <https://doi.org/10.4324/9780203852187>
- Spradley, J. (1980). *Participant observation*. Holt, Rinehart and Winston.

- Stockleben, B., Thayne, M., Jäminki, S., Haukijärvi, I., Mavengere, N. B., Demirbilek, M., & Ruohonen, M. (2017). Towards a framework for creative online collaboration: A research on challenges and context. *Education and Information Technologies*, 22(2), 575-597.
- Schofield, J. (2018, April 19). What's the best email service that doesn't scan emails for ad-targeting? *The Guardian*.
<https://www.theguardian.com/technology/askjack/2018/apr/19/whats-the-best-email-service-that-doesnt-scan-emails-for-ad-targeting>
- Walker, N., Hopper, T., Fu, H., & Sanford, K. (2017). A roadmap to digital ePortfolio proficiency: Creating the conditions for cross-course, cross-program and cross-life professional learning. *The AAEEBL EPortfolio Review*, 1(2), 58–72.
- Warrell, J., & Jacobsen, M. (2014). Internet research ethics and the policy gap for ethical practice in online research settings. *Canadian Journal of Higher Education*, 44(1), 22-37.
- Internet message access protocol. (2021, January 28). In *Wikipedia*.
https://en.wikipedia.org/wiki/Internet_Message_Access_Protocol

Author Note

Tim Hopper (Ph.D., University of Victoria) is a Professor in the School of Exercise Science, Physical and Health Education. Please direct correspondence to thopper@uvic.ca.

Hong Fu (Ph.D., University of Victoria) is a Research Associate and Sessional Instructor in the Department of Curriculum and Instruction. Please direct correspondence to fuhong@uvic.ca.

Kathy Sanford (Ph.D., University of Victoria) is a Professor and Interim Graduate Advisor in the Department of Curriculum and Instruction. Please direct correspondence to ksanford@uvic.ca.

Thiago Alonso Hinkel (M.A., University of Victoria) is a doctoral student in the Department of Curriculum and Instruction. Please direct correspondence to thiagohinkel@uvic.ca.

Copyright 2021: Tim Hopper, Hong Fu, Kathy Sanford, Thiago Alonso Hinkel, and Nova Southeastern University.

Article Citation

Hopper, T., Fu, H., Sanford, K., & Hinkel, T. A. (2021). YouTube for transcribing and Google Drive for collaborative coding: Cost-effective tools for collecting and analyzing interview data. *The Qualitative Report*, 26(3), 861-873. <https://doi.org/10.46743/2160-3715/2021.4639>
