CAIS Paper: Towards Understanding of Human Rights Researchers' Data Analysis Practices - An Interview Study

Lu Xiao (the University of Western Ontario) and Jillian R. Kavanaugh (Boston Children's Hospital)

Abstract: The impetus to assist human rights researchers in data analysis is stronger than ever. However, little is known in the literature on human rights researchers' practices in collecting, managing, and analyzing their research data. Addressing this gap, we interviewed 13 researchers whose research areas are related to human rights issues.

1. INTRODUCTION

Human rights research is a multidisciplinary field in which researchers, whether in academic institutions or independent non-profit organizations, collect and analyze data from a variety of geographical areas - and oftentimes the volume of data can be overwhelming. Human rights researchers use a wide range of analysis techniques from statistical analysis and modeling to examining correlations among different factors of human rights violations, and to studying the impact of human rights violation events ethnographically. Researchers also interact with relevant data in various sorts of formats (e.g., textual, audio, and video) and conduct both qualitative and quantitative content analyses. These data are increasingly accessible as Internet-based resources, such as the Center for Human Rights Documentation and Research at Columbia University, the Human Rights Documentation Initiative (HRDI) at the University of Texas at Austin, and the University of Connecticut's Human Rights Film Collection. These online resources exist to digitally preserve and share crucial human rights documentation with other researchers.

While these trends in the generation and access of human rights violation data offer more data analysis opportunities to human rights researchers, they also present challenges in the qualitative data analysis because of the ever-growing size of the data resources. Computer scientists and digital humanities scholars have started to explore computational approaches to address these challenges (e.g., Miller, Li, Shrestha, and Umapathy, 2013; Strata, 2013). However, there is limited understanding about human rights researchers' data analysis practices. We have interviewed 13 human rights researchers to understand their data analysis practices in human rights research; their usage of computer programs in their analysis processes; and their expectations of such a qualitative data analysis tool.

2. RELATED WORK

There are some studies about qualitative data analysis processes and the important factors. Frost et al. (2010) pointed out that it is important to know how qualitative researchers' expectations and assumptions are understood, applied and presented, and to make them transparent to understand the process of transforming data in qualitative research. The authors also suggested the use of different analysis techniques in qualitative analysis to complement each other to

overcome shortcomings. To understand the meanings attached to qualitative research practice and the perceived challenges posed by contemporary innovations in data management, access, and analysis through electronic archives, Broom, Cheshire and Emmison (2009) conducted six focus groups with 37 Australian qualitative researchers. Their results reveal key issues such as the idea that the researcher has a special relationship with the data which prevents anyone else from analyzing them in their original context and which leaves the data 'disembodied' when archived, as well as concerns over research ethics and data ownership.

To help novice qualitative researchers understand what tool to use for which stage of analysis, Peters and Wester (2007) emphasized the importance of providing specific and detailed instructions to link computer tools to research methodology in order to help researchers better understand the tools and support their research.

3. RESEARCH METHODS

Our interview questions were divided into four sections. The first section was about the interviewee's research background, and the data that he/she has worked with in their research activities. For example, we asked interviewees to describe the data they often analyze, such as the format (photos, text, video, etc.), the size, the type (primary or secondary data), and the source (publicly available or private). In the second section, we asked the interviewees to think of a concrete example of data analysis in their projects and describe the analysis process. We also asked them to provide details like what they were looking for in the data, whether they used software programs for analysis (and if so, the name of the programs), and the software program features that they found most useful in the analysis. Interested in how researchers coordinate on shared data during the analysis process, we asked our interviewees how many people are generally involved in data analysis projects, and if there are indeed multiple analysts, we asked how they coordinate with each other in the process of analyzing data and integrating the results. In the last section, we sought their expectations on a software program that would support qualitative analysis of human rights data, either small or large scale. We asked for concerns or issues that should be addressed, three most important design features desired in such a program, and features they considered ineffective and should be excluded (based on their experiences).

Through purposive and snowball sampling techniques, we interviewed 13 North American academic and non-academic researchers of human rights violation research. While our 13 interviewees were based in North America, the data they analyzed originated from a variety of countries spanning several continents. All interviews were conducted in a semi-structured format. The interviews were 60 – 90 minutes long except one interview which lasted 25 minutes because the interviewee's background and experiences are different from what we were looking for. One interview was via phone, one via Skype, and the rest were conducted face-to-face at locations of the interviewees' choices (e.g., their offices). The first author and a research assistant were both present in the interviews taking the roles of asking questions and taking notes. All interviews were audio recorded.

To analyze the interview data, we identified salient themes in individual transcripts and searched for repetitions within and across narratives and field notes (Ryan & Bernard, 2003).

4. FINDINGS
4.1. Interviewees' Background and Data Analysis Practices
The interviewees were from independent human rights organizations or universities, and their academic background covered eight disciplines. They've worked on human rights research issues from a few years to over 35 years. All but one researcher have dealt with primary data, with some being publicly accessible, some required special access. These data varied in formats and include text, number, audio/video, and artifact (e.g., images, museum objects). They came from various sources internationally, such as police records, court decisions, U.S. Department of State Human Rights Reports, and interviews conducted by the researchers themselves.

The interviewees reported the use of quantitative and qualitative analysis techniques, with majority of the interviewees focused on the qualitative analysis approach. The size of the data in the qualitative analysis is within hundreds of pages (e.g., 40 interviews, documents of 20 countries, etc). The researchers acknowledged the lengthy process of reading, annotating and analyzing the textual data. Interestingly, no interviewee used any qualitative analysis software program, whether a comprehensive program like QSR International's NVivo, or a human-rights specific program like HURIDOCS' OpenEvsys in the process. Instead, they either annotated the textual documents using paper and pencil technique, or took notes when analyzing the audio/video recordings. In fact, many interviewees were not even familiar with qualitative software programs. Although the datasets could be large (e.g., in one project the researchers had to read all the countries' reports and judge the countries' human rights violation situation based on the reading), all interviewees reported using traditional and low technical methods to analyze textual data, such as reading through PDFs, Word documents, or working almost exclusively with print documents. One interviewee explained that he/she used Word processing and spreadsheet software to store excerpts that had been extracted from the textual documents.

Although qualitative analysis software was not used, several interviewees used collaborative software programs such as Dropbox and Basecamp to aid the sharing and management of their data or documents for data analysis. An interviewee whose work was done mainly through ethnographic studies explained that he/she used Facebook to communicate with people of the studied countries and the analytical process was a collaborative interpretation process.

As well, several interviewees created their own databases for storing and accessing their documents, and one interviewee created their own open source software program for researchers to share audio/video interviews within the research team.

4.2. Expectation of a New Software Program for Qualitative Analysis
With regards to the expectations of qualitative data analysis software, the interviewees presented a wide range of characteristics that would be useful for their research purposes. Two interviewees would like a program that would recognize patterns and meaning in the text being analyzed. One suggested, "create a more effective, searchable database with all of the documents held by the Inter-American Court, and again, that would prompt the user to look for certain types of patterns and certain types of information that could be organic, be changing. I wouldn't want it to just be a Google search of a bunch of documents." When asked to define a "pattern," the

interviewee provided an example of how the court's decisions often do not mention violence against women. For example, certain cases "may in fact be cases of violence against women, or they may in fact contain within them a story about patterns of violence against women that are not being treated by the court as such, so the court is talking about this individual being disappeared or murdered". Another interviewee was also interested in a program that would locate patterns, stating that such a program would "allow me to find patterns that I wouldn't be able to find otherwise, and also quickly, which is great, and of large interview collections".

Interestingly, one interviewee would like to see a program that not only assists in the analysis of data, but also the management of data, in order to assist in the collaborative efforts of qualitative data analysis. The interviewee, who supervises students analyzing U.S. Department of State Human Rights Reports, stated, "one of the things we struggled with was how to have the students manage the massive amounts of data…and be able to find things quickly." Other requested features included offering multi-language support, reducing the workload of collecting disaggregated survey data, enabling transcription of the audio files, and enabling adding and searching detailed metadata. One interviewee also suggested that concordance software features are desired. We noted that these expected features are related to data handling and management issues.

5. CONCLUSION AND FUTURE WORK

In this paper, we report an interview study that helped us better understand the human rights researchers' data analysis practices, the characteristics of their research data, and their expectations on qualitative data analysis software programs. We found that the use of qualitative software programs was uncommon even though all the interviewed researchers have dealt with the situation of analyzing large amount of textual data in the projects. However, collaborative software programs have been used among the researchers to coordinate on sharing the data or intermediate analysis results (e.g., coding results, reflection blogs). Despite of little or no experiences with data analysis software for qualitative research, our interviewees knew what they would want in such a tool; for example, desired features included an effective and searchable database, pattern and meaning recognition from text, multilingual support, annotation features, audio/visual archiving, and general data management. Altogether, these suggestions were valuable for brainstorming a tool that would be suitable for researchers in this area. Additionally, by examining how researchers in the complex and multidisciplinary field of human rights research typically analyze qualitative data, we hope to create a call for action towards further understanding of how custom software can assist these researchers in this field of global importance.

We are currently conducting an online questionnaire study, based on the interview results. Our purpose is to reach out more human rights researchers to test the extent of generalizability of our findings.

REFERENCES:

Broom, A., Cheshire, L., & Emmison, M. (2009). Qualitative researchers' understandings of their Practice and the Implications for Data Archiving and Sharing. Sociology, 43(6), 1163-1180.

Frost, N., Noals, S.M., Brooks-Gordon, B., Esin, C., Holt, A., Mehdizadeh, L., & Shinebourne, P. (2010). Pluralism in qualitative research: the impact of different researchers and qualitative approaches on the analysis of qualitative data. Qualitative Research, 10(4), 1-20.

Miller, B., Li, F. X., Shrestha, A., & Umapathy, K. (2013), Digging into Human Rights Violations: phrase mining and trigram visualization, Digital Humanities'13 (July 16 – 19, Lincoln, NE, USA)

Peters, V. & Wester, F. (2007). How qualitative data analysis software may support the qualitative analysis process. Quality & Quantity, 41, 635-659.

Ryan, G. W., & Bernard, H. R. (2003). Techniques to identify themes. Field Methods, 15(1), 85-109

Strata, O., (2013). Machine Learning for Human Rights Data Science for Social Good fellows partner with Ushahidi, accessed on Dec. 8, 2013, http://strata.oreilly.com/2013/07/machine-learning-for-human-rights.html