

# Literature Review: How Content Analysis Is Used to Study Human Rights Violations in Political Science?

**Tatiana Vashchilko**

Faculty of Information and Media Studies  
University of Western Ontario  
[tvashchi@uwo.ca](mailto:tvashchi@uwo.ca)

**Lu Xiao**

Faculty of Information and Media Studies  
University of Western Ontario  
[lxiao24@uwo.ca](mailto:lxiao24@uwo.ca)

---

## Abstract

Our study focuses on the review of the existing datasets on human rights violations to understand the types of data sources used to construct the datasets and the methodologies to measure the degree of human rights violations in different countries. This poster paper reports, preliminary findings from Political Science research in 1971-2012 on the scope of employed human right violations' data and methodologies.

*Keywords:* human rights, content analysis, political science

---

## Introduction

Quantitative analysis of human rights violations is important for systematically testing existing theories explaining human rights violations along with their consequences, and design of effective policy instruments to prevent future human rights violations. Amnesty International and The US Department of State's Human Rights Reports were the first ones in the early 1980s to provide any systematic textual description of human rights violations across different countries on an annual basis. Later on, these reports have become the basis for two-three large-scale datasets with the quantitative measures of the degree of human rights violations across multiple dimensions (e.g., civil rights, political rights). As the importance of reporting human rights violations by governmental and non-governmental organizations rose, the primary sources with the rich textual data offering additional details about the victims, perpetrators, and events have emerged. With the recent availability of primary information on human rights violations, and advances in natural language processing (NLP) to quickly identify and extract systematic data from textual information, this paper analyzes how scholarly community have been utilizing evidences of human rights violations. The paper contributes to research on human rights violations by systematically accounting for the existing use of primary and secondary data on human right violations in Political Science, one of the main disciplines studying this phenomenon. In this poster paper, we report preliminary findings on the identification of the majority of human rights data sources employed in Political Science based on the analysis of peer-reviewed articles in 1971-2012. The paper deliberates first on the corpus, then proceeds with the description of the coding scheme, and finally discusses the results and implications.

## Corpus and Coding Schema

There are two major databases in political science research: Worldwide Political Science Abstracts (WPSA) operated by ProQuest and International Political Science Abstracts (IPSA) operated by Ovid Wolters Kluwer. We chose WPSA in our study because it covers larger period of publication time (1923 – present), and offers greater journals' coverage. For example, searching "human rights" phrase in IPSA produced only 5,412 results, peer-reviewed publications, whereas WPSA produced 21,949. Because our goal is to understand the use of human rights violation data in the research studies and how they were analyzed, we used filters such as "data\*", "analysis," "case stud.," "empiric\*," or "statistic\*" to further specifying the literature body we are interested in. With only peer-reviewed English articles included, our search produced 2,796 items published in 1971-2012 (as of July 30th, 2012). We then used two criteria to

select the papers for further analysis. The first and most important one is that the reported study should have used evidence on human right violations in the analysis (e.g., testimonies of victims, photographs, etc). The second criterion intended for including each of the quantitative datasets on human rights violations only once in our analysis papers and, therefore, had two requirements: 1) the quantitative datasets of human rights violations had to be constructed by the authors of a peer-reviewed article under consideration, and 2) the measures of human rights violations in the quantitative datasets had to be based on textual or non-textual human rights violation documents. Including only these types of peer-reviewed publications advanced the understanding of content analysis used to extract quantitative characteristics of human rights violations from textual and non-textual sources. After these criteria were applied, an expert opinion of a political scientist, one of the authors of this paper with a doctorate degree in Political Science, was used to determine whether a paper should be included or not based on these criteria by reading through the titles, abstracts, and sometimes full texts of the papers, if needed, of all selected 2796 peer-reviewed articles.. After this first step, 152 potential articles from Political Science journals were selected for the subsequent more nuanced analysis. Finally, after careful consideration of 152 articles, 31 peer-reviewed articles published in Political Science were included in the corpus to conduct the analysis of the utilized sources on human rights violations. To make sure that our selection is comprehensive, we compared our generated list with the reference lists of major literature review papers such as Green (2001) and Milner et al., (1999). After the extensive literature search, careful identification and subsequent verification processes, the belief is that the corpus includes the vast majority of peer-reviewed articles in Political Science employing at least one information source on human rights violations

We coded the selected 31 articles focusing on the information sources and data analysis. The example codes include characteristics of the sample (e.g., sample size, year coverage), data source, data accessibility, data format, and the kinds of data analysis methods. We also coded the software programs used in the analysis, if any.

## Preliminary Findings

We found that the papers in the sample were predominantly published in the last five years (Figure 1). Apart from the growing scholarly interest in the issues of human rights abuses, there seems to be two main reasons that drive this trend: 1) the availability of information on human rights violations, and 2) the advances in software development for content analysis. For example, the need for collection and preservation of information on human rights violations has been brought to light with the potential of better identification of victims and predators of past crimes if employing systematic qualitative and quantitative analysis (Rosenberg, 2012).

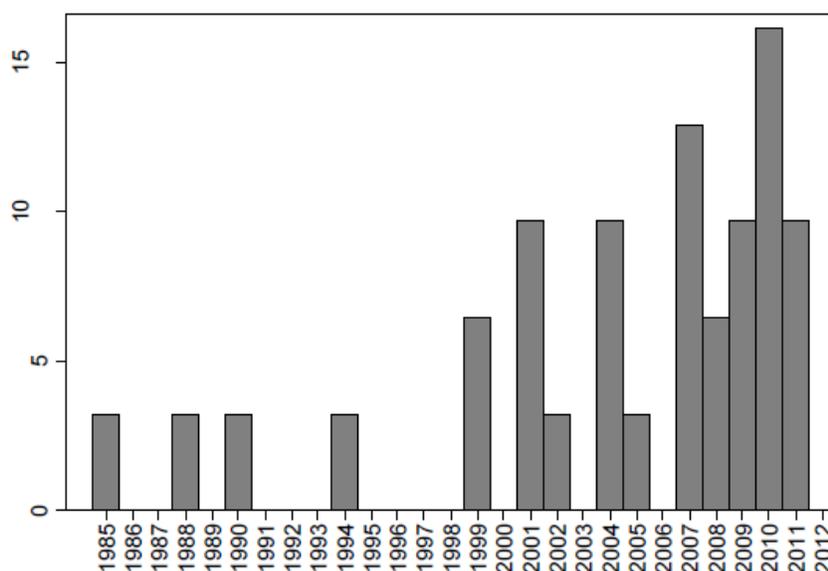


Figure 1. Distribution of Papers in Political Science by the Year of Publication

In terms of qualitative analysis, we considered three modes: *manual*, *semi-automated*, and *automated*. Both the hand coding and the use of software programs like NVivo are considered manual coding. Two studies used algorithms that take coding schema and corpus as input and generated coded results as the output. For example, Chapman (2007) uses semi-automated analysis to evaluate the experience of the South African Truth and Reconciliation Commission (TRC) and assess whether the TRC was successful in developing a model to promote intergroup forgiveness and reconciliation. In particular, to understand the full range of victimization experience, the transcripts of 429 victims were quantitatively coded using ATLAS/ti computer-based program. No study used completely automated approach for content analysis, that is, the use of an algorithm that takes corpus as the input and generates coded results as output with natural language processing and text mining mechanisms.

Most of the papers were using publicly available sources of information (Figure 2). Also, 17 studies relied on secondary data, 12 on primary data, and 2 on both. We found that if sample included many countries and time periods the studies tend to rely on secondary and public data. The “two most widely distributed and read sources on information on countries’ human rights practices” are the US Department of State Country Reports and the Amnesty International annual reports (Poe, Carey, & Vazquez, 2001, p. 653). Most of the papers on human rights violations used the event-level or country-year level of analysis (Davenport & Ball, 2002).

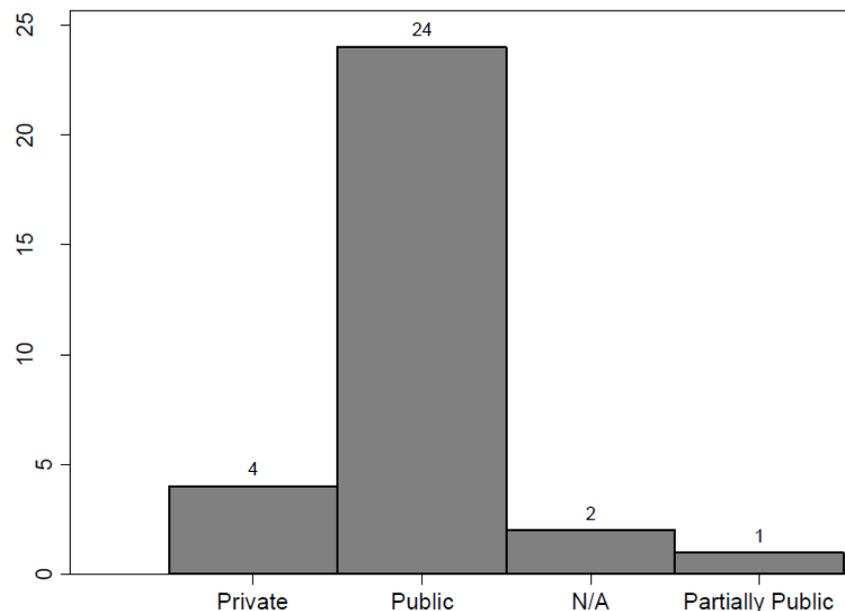


Figure 2. Use of Public and Private Information in Political Science Research on Human Rights Violations

The majority of the papers, which used primary data and private information, focused on one country or two countries to investigate predominantly the specifics of human rights violations committed in those countries. However, in general, very little primary documentation on human rights violations is available in research community. Montgomery (2001) pointed out four reasons that contributed to this: 1) NGO’s concern about potential access to the victims’ testimonies by the perpetrators; 2) inability of libraries and archives to recognize the importance of preserving human rights violations’ documentation; 3) financial constraints of existing archives of human rights violations documents; and 4) reluctance of private foundations to fund the human rights violations archives.

## Discussion and Future Work

Political science developed two main approaches to measure human rights violations. One way to measure human rights violations is by means of coding cases of repressive events from media, so-called events-based approach (Milner, Poe, & Leblang, 1999). Another one is standards-based approach, which measures the degree of human rights violations in different countries by coding governments’ reports on human rights and categorizing countries according to a “predetermined criteria” (Milner et al. 1999, p.

421). For example, a very well-known measure of human rights violations, Political Terror Scale (PTS), was developed in the 1980s (Stohl & Carleton, 1985), and has been used by many researchers in Political Science to code various textual reports on human rights violations across different countries (Gibney, Cornett & Wood 2012).

The standards-based approach is the preferred one in Political Science due to lesser extent of potential biases and greater comparability of degree of human rights violations across countries. However, it sacrifices the details of the human rights violation events. Sylva and Wyly (2001) demonstrated the emerging benefits of the events-based approach Investigating the nature of coverage of Sierra Leone civil war in the Western media, specifically in New York Times, Sylva and Wyly (2001) identified that “the single most frequent reference in the articles analyzed was to the types of atrocities committed or alleged. Almost half provided information on human rights violations, and more than one-third included estimates of the number of war casualties” (p. 45). Thus, Sylva and Wyly (2001) demonstrated that a major western newspaper reported on human rights violations in a detailed manner. With globalization and the spread of the Internet, the events-based approach based on the media coverage of various military conflicts might generate improved quantitative human rights violations indicators, which can, if not supersede, at least be comparable to the standard-based approach and depending on the aggregation of the media outlets, less biased than the standard-based approach. This implies the need of developing algorithms for analyzing the amount of data introduced through such approach. It is also expected that the availability of such tools for analyzing big data may affect how human rights violation studies are conducted in the future. The development of such algorithms is the ultimate goal of this funded project (funding agency: Social Sciences and Humanities Research Council in Canada).

## References

- Carleton, D., & Stohl, M. (1985). The Foreign Policy of Human Rights: Rhetoric and Reality from Jimmy Carter to Ronald Reagan. *Human Rights Quarterly*, 7(2), 205-205.
- Chapman, A. R. (2007). Truth Commissions and Intergroup Forgiveness: The Case of the South African Truth and Reconciliation Commission. *Peace and Conflict: Journal of Peace Psychology*, 13(1), 51-69. doi:10.1080/10781910701226501
- Davenport, C., & Ball, P. (2002). Views to a Kill. Exploring the Implications of Source Selection in the Case of Guatemalan State Terror, 1977-1995. *Journal of Conflict Resolution*, 46(3), 427-450. doi: 10.1177/0022002702046003005
- Gibney, M., Cornett, L., & Wood, R. (2012). *Political Terror Scale 1976-2008*. Retrieved from: <http://politicalterrorsscale.org/>
- Green, M. (2001). What We Talk about When We Talk about Indicators: Current Approaches to Human Rights Measurement. *Human Rights Quarterly*, 23(4), 1062-1097.
- Milner, W. T., Poe, S. C., & Leblang, D. (1999). Security Rights, Subsistence Rights, and Liberties: A Theoretical Survey of the Empirical Landscape. *Human Rights Quarterly*, 21(2), 403-403.
- Montgomery, B. P. (2001). Human Rights: A Survey of Archival Sources in the United States and Canada. *Human Rights Quarterly*, 23(2), 431-463.
- Poe, S. C., Carey, S. C., & Vazquez, T.C. (2001). How Are These Pictures Different? A Quantitative Comparison of the US State Department and Amnesty International Human Rights Reports, 1976-1995. *Human Rights Quarterly*, 23(3), 650-677.
- Rosenberg, T. (2012). The Body Counter. *Foreign Policy*(192), 88-96.
- Silva, J. A., & Wyly, E. K. (2001). Between Africa and the Abyss: Globalization, Media, and the Invisibility of a Continent. *Geographical Bulletin*, 43(1), 36-46.