

Classifying the Content of Physical Integrity Rights Allegations using Machine Learning Methods*

Rebecca Cordell[†]

K. Chad Clay[‡]

Christopher J. Fariss[§]

Reed M. Wood[¶]

Thorin M. Wright^{||}

Abstract

We present and evaluate an automated coding process that produces categorical information about physical integrity rights violations described in annual human rights reports. The coding process builds on an existing corpus of human rights allegations. Allegations are sentences taken from country-year human rights reports published annually by Amnesty International, Human Rights Watch, and the US State Department. Each allegation sentence presents information about a specific type of physical integrity violation (disappearances, torture, killing and political imprisonment) and may also contain information about the date, actor type, location, scope, intensity, range, victim type, and whether or not the violator of the right was held accountable. The information content of the allegations varies, and each allegation only contains information about a subset of these items. Therefore, the information for each allegation is based on a probability model that relates the text contained in each allegation to human coded categorical data of each item. The coding of each allegation sentence is quite costly in terms of human coding time. Our human identified allegation data covers a sample of countries: Angola, Belarus, India, Mexico, Nigeria, Philippines, and the United Kingdom for every odd year report between 1999 and 2016. We use this training data to code each repression allegation for all countries in the world from 1999-2016 using a supervised machine learning approach. We present support vector machine, naïve bayes, logistic regression based probabilities and use cross validation to assess model performance.

* This project receives funding from the National Science Foundation (award number 1627464, 1626775, 16267101). For more information, please visit: <http://snarpdata.org/>.

[†] Postdoctoral Research Associate, School of Politics and Global Studies, Arizona State University, Rebecca.Cordell@asu.edu.

[‡] Associate Professor, Department of International Affairs, University of Georgia, kcclay@uga.edu.

[§] Assistant Professor, Department of Political Science, University of Michigan, cjfariss@umich.edu.

[¶] Associate Professor, School of Politics and Global Studies, Arizona State University, Reed.Wood@asu.edu.

^{||} Associate Professor, School of Politics and Global Studies, Arizona State University, Thorin.Wright@asu.edu.