

Identifying Human Rights Abuses



www.business-humanrights.org

The Problem

- business-humanrights.org is looking for help identifying reports of human rights abuses involving companies.
- **Examples:** child and forced labor in manufacturing, miners throwing farmers off their land.
- Potential sources of new violations:
 - Newspaper articles.
 - Tweets.
 - Other social-media postings (harder).

Data Available

- B-HR.org will provide a scraping of all the articles they have identified as relevant to a corporate human rights abuse
 - Classified by general area of the problem.
 - Website currently has 86,000 links.
- Stanford has:
 - Twitter sample that can be used for experiments in identifying relevant tweets.
 - News articles for the same purpose.

Challenges (Need not address all)

- Model the relevance of words used in identified news articles to detect similar articles.
 - Harder: identify relevant tweets.
- How would you use the Twitter interface to identify good candidate tweets that might be relevant, without missing too many?
- How would you mine other social media (nonstreaming) to discover potential human-rights violations?