

Toward Automating Detection of Human Values in the Nuclear Power Debate

Emi Ishita, Toru Oga

Kyushu University, Japan
Ishita.emi.982@m.kyushu-u.ac.jp,
toga@law.kyushu-u.ac.jp

An-Shou Cheng

National Sun Yat-sen University, Taiwan
ascheng@mail.nsysu.edu.tw

Kenneth R. Fleischmann

The University of Texas at Austin, USA
kfleisch@ischool.utexas.edu

Takayama Yasuhiro

National Institute of Technology
Tokuyama College, Japan
takayama@tokuyama.ac.jp

Douglas W. Oard

University of Maryland, USA
oard@umd.edu

Yoichi Tomiura

Kyushu University, Japan
tom@inf.kyushu-u.ac.jp

ABSTRACT

This paper describes the creation of a corpus of newspaper articles about the Fukushima nuclear disaster, a coding frame for content analysis of human values expressed or reflected in that corpus, and preliminary results for automation of the content analysis. Understanding the human values that motivate sentiment towards an idea can help to characterize the basis for that sentiment, and this work is a first step toward applying that approach to positions on controversial events reported in the news.

Keywords

Human values, test collection, automated content analysis.

INTRODUCTION

Newspaper articles provide a lens for viewing prominent events within their social, political, economic and cultural contexts. For controversial issues, journalists typically seek to report on a range of perspectives, along with information about the basis for those perspectives. Our goal is to study how human values can help to explain positions that people take on controversial topics (Fleischmann, 2014), and thus we are interested in coding news stories with the human values that they reflect. Because the scale of modern news media content exceeds what any person could code by hand, we are also interested in exploring the potential for automating this process using supervised machine learning.

Our approach is inspired by our earlier work (Takayama et al., 2014), in which we developed a Latent Value Model (LVM) to automatically code a small set of human values in sentences from about 100 statements prepared for legislative and regulatory bodies in the Net Neutrality

debate in the United States. The automatically coded human values were then compared with manually assigned human value labels (Cheng et al., 2012) to compute classification accuracy measures such as precision and recall, finding that the LVM classifier (which operates on both single words and pairs of words) was more accurate than several widely used classifiers for this task. Our work differs from that earlier work in several ways: (1) we focus on third-person news reporting rather than first-person prepared statements, (2) we code entire articles rather than individual sentences, and (3) we code for a broader range of human values.

In the remainder of this paper, we introduce our annotated news article corpus, we report some initial experiment results using our LVM classifier, and we look ahead to the next steps in our work-in-progress.

NEWS ARTICLE CORPUS

Our focus is on the debate over the future of nuclear power in Japan that resulted from the Fukushima Daiichi nuclear reactor disaster on March 11, 2011. The Great East Japan Earthquake that occurred on that date damaged the reactor, leading to an immediate evacuation, and ultimately leading to the temporary shutdown of every nuclear power generation facility in the country. This was a classic example of what Taleb (2007) has called a “black swan” – an unanticipated event with the power to change perceptions regarding what is possible. The controversial decision to shut down all 54 of Japan’s nuclear reactors, which had previously supplied 30% of the nation’s power, had substantial consequences both for economic activity and for the activities of daily life in this highly industrialized society. The subsequent decision to restart some of those reactors, and the government’s ultimate decision to continue to rely on nuclear power for energy (in contrast, for example, to the contemporaneous German decision to permanently end nuclear power generation) has proven even more contentious.

ASIST 2017, Washington, DC | Oct. 27-Nov 1, 2017

Authors Retain Copyright.

Value	2011	2012	2013	P	R	K
Human welfare	237	255	203	0.38	0.72	0.27
Importance	176	185	172	0.27	0.63	0.08
Effectiveness	128	86	87	0.15	0.62	0.09
Power	95	86	112	0.15	0.63	0.10
Personal welfare	63	74	91	0.13	0.68	0.09
Nature	67	52	95	0.12	0.67	0.28
Other	61	45	75	0.07	0.49	0.11
Law and order	49	71	47	0.08	0.55	0.34
Wealth	55	50	43	0.09	0.73	0.27
Innovation	31	35	61	0.07	0.67	0.26
Independence	24	31	24	0.05	0.70	0.13

Table 1. Article counts; LVM Precision & Recall; Kappa.

The time scale over which these controversies unfolded calls for the development of a new test collection. For our initial work, we have therefore collected 13,821 news articles in Japanese from the Japanese daily newspaper CD Mainichi Shimbun (Mainichi Newspaper Annual Full-Text Data Base) between 2011 and 2013. We selected the articles using the query "[原発 (nuclear power plant) OR 原子力 (nuclear power)] AND [福島 (Fukushima)]" in the heading or body of the article. This resulted in 6,629 articles from 2011, 4,036 articles from 2012, and 3,156 articles from 2013. We then randomly selected 700 articles from each year for manual coding, resulting in the 2,100 manually coded articles that we used with 10-fold cross-validation for training and testing in these experiments.

Coding Frame and Coding

Trained coders assigned six types of labels to each of the 2,100 articles. For four of those types (Type, Phase, Source, and Attitude), the coders assigned a label to each article. Any number of labels could be assigned for the Media and Values types. In this paper, we focus on the Values labels. Our Values coding frame includes 11 human values (see Table 1). One author, a social scientist, initially built the Values coding frame to analyze the contemporaneous discussion of nuclear power policy in Taiwan using a collection of articles from a Taiwanese newspaper. We first prepared a coding manual and then hired and trained coders. The coders examined the headline and body of each story and assigned Values labels. Two coders independently coded each article. Inter-coder agreement scores computed using Cohen's (1960) Kappa (K) are shown in Table 1.

PRELIMINARY EXPERIMENTS

We conducted experiments with two types of classifiers (LVM, and SVM^{light} Support Vector Machines with several kernel functions). In both cases, we first used the standard Juman package to perform tokenization and morphological analysis for Japanese and then selected noun, verb, adjective, nominal prefix and suffix, and numerative parts of speech as the term space for our classifiers. Articles

included headlines (average of 12.7 terms), and bodies (average of 271.3 terms). Table 1 shows our best overall results for precision and recall, which were obtained using LVM. For example, for Human welfare, LVM achieved a precision of 0.38 and a recall of 0.72, whereas our best SVM (linear kernel) achieved a precision of 0.54 and a recall of 0.23. We note that LVM consistently yields higher recall than precision, whereas for SVM the opposite is true. These results are comparable to those reported in Takayama et al. (2014) for the value Honor, the one category in those experiments that had a comparable number (317) of positive training examples.

NEXT STEPS

These results will serve as a baseline for comparing future improvements. Our next step must be to refine our coding process to improve inter-coder agreement; in other work, we have found that asking annotators to focus on the lead paragraph of a news story can help to improve consistency. Next, we will try semi-supervised classification techniques that can benefit from the large number of unlabeled articles, and with joint models that can assign multiple types of labels simultaneously. It is important to consider how best to tune the tradeoff between precision and recall to optimally support social science research at Web scale, which can guide future work with committees of diverse classifiers to achieve that balance. Ultimately, we hope to unify our test collection with a collection on the same topic developed in Taiwan for Chinese-language news; a first step toward creation of a multilingual collection for studying the nuclear power debate. Throughout this process, we will share our coding manual and coding results publicly so that others can replicate our work and extend it in novel ways.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number JP25280118.

REFERENCES

- Cheng, A.-S., Fleischmann, K. R., Wang, P., Ishita, E., & Oard, D. W. (2012). The role of innovation and wealth in the Net neutrality debate: A content analysis of human values in Congressional and FCC hearings. *JASIST*, 63, 1360-1373.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurements*, 20, 37-46.
- Fleischmann, K. R. (2014). *Information and Human Values*. San Rafael, CA: Morgan & Claypool.
- Takayama, Y., Tomiura, Y., Ishita, E., Oard, D. W., Fleischmann, K. R., & Cheng, A.-S. (2014). A word-scale probabilistic latent variable model for detecting human values. *CIKM*, 1489-1498. Shanghai.
- Taleb, N. (2007). *The black swan: The impact of the highly improbable*. New York: Random House.