

The Deliberative E-Rulemaking Project (DeER): Improving Federal Agency Rulemaking Via Natural Language Processing and Citizen Dialogue

Peter Muhlberger
Texas Tech University
Dept. of Political Science
116 Holden Hall Boston & Akron
Lubbock, Texas 79409
001-806-791-7723
pmuhl830@gmail.com

Nick Webb
University at Albany, SUNY
Institute of Informatics, Logics and
Security Studies
Albany, NY, 12222
001-518-442-3082
nwebb@albany.edu

Jennifer Stromer-Galley
University at Albany, SUNY
Dept. of Communication, SS340
Albany, NY 12222
001-518-442-4873
jstromer@albany.edu

ABSTRACT

Many scholars believe that electronic rulemaking has great but largely untapped potential to expand the public's democratic input and improve federal agency regulatory rules. The existing federal rulemaking process, however, elicits many redundant and poorly considered comments, and some participants challenge the legitimacy of agency rulemaking decisions. This Deliberative E-Rulemaking (DeER) project tests the application of Natural Language Processing (NLP) technologies and public deliberation techniques to improve the quality and organization of public comments and the legitimacy of the federal rulemaking process. This paper describes the goals, innovations, research design, technology, and preliminary results of the DeER project. Pilot results suggest a number of ways to improve future phases and indicate improvements in participants' sophistication in thinking about government. The social science herein advances research on measuring the quality of public deliberation and the effects of deliberation on individuals and communities.

Categories and Subject Descriptors

K.4.0 [Computers and Society: General]

General Terms

Management, Performance, Experimentation, Human Factors.

Keywords

Rulemaking, Natural Language Processing, Interactive Question Answer, Dialogue Analysis, Summarization, Democratic Deliberation, Deliberation Quality, Distributed Intelligence

1.Goals and Innovations

The primary objective of this project is to expand the value and legitimacy of electronic rulemaking while using this practical context to advance Natural Language Processing (NLP) and the social science of deliberative groups. By leveraging both NLP and a deliberation technique called Multi-Level Deliberation (MLD), DeER seeks to address several key problems relating to e-

rulemaking: 1) problems of scale such as output volume and information sharing across large numbers of participants and 2) poorly informed and unsophisticated participants.

Our objective for the NLP component of the project is to combine such NLP technologies as Interactive Question Answering (QA), Dialogue Analysis, and Summarization into a viable Discussion Facilitation Agent (DiFA) that will keep users informed about changes and developments in the deliberation and summarize key arguments. A second and complementary key component of this effort will be Multi-Level Deliberation (MLD). In such deliberation, small groups will discuss the rulemaking topic and then, through proportional representation voting, select two members that will represent the group in a higher-level group that represents multiple lower-level groups. Thousands of participants can be represented in a few levels. More informed and engaged participants and better ideas should flow toward higher-level groups. Information can also flow back to enhance the legitimacy of final agency decisions. Interaction among levels can promote the emergence of distributed intelligence. Agency officials can directly interact with participants in the highest-level groups.

2.Research Design

The project will involve four phases of experimental research. The first two phases test the NLP technologies and social science measurement approaches on small samples college students. The last two phases will involve large samples of actual public participants in agency rulemakings. The final 3X2 experiment crosses MLD, non-MLD deliberation, and standard, non-deliberative rulemaking participation with the presence or absence of DiFA. The success of the various conditions of these experiments will be determined using survey and focus group measures of agency official and participant perceptions and evaluations; a content analysis measure of the cognitive sophistication of comments; both human-coded and automated content analyses of the quality of deliberation; measures of the impact of the deliberations on participants, including knowledge, trust, citizenship, and communicative rationality; perceptions of the legitimacy of decisions; DiFA usage patterns; and continued

participation.

3.Key Technology

In the initial phases, we concentrate on NLP technologies that can inform users before and during their deliberation regarding their deliberation topic. We have integrated a discussion board mechanism, the discussion module of the Drupal content management system, with our existing Interactive QA system, HITIQA. The HITIQA system allows users to ask unconstrained natural language questions over a data space and returns short snippets or paragraphs in which the answer can be found. For example, the question “*Why is net neutrality important?*” returns a first paragraph that contains the following:

“When asked how important a ‘Consumer Bill of Rights’ would be that prevented Internet Service Providers from blocking or degrading access to Internet sites and services, 78 percent indicated that such a bill would be important, with 59 percent of that group calling it ‘very important.’....”

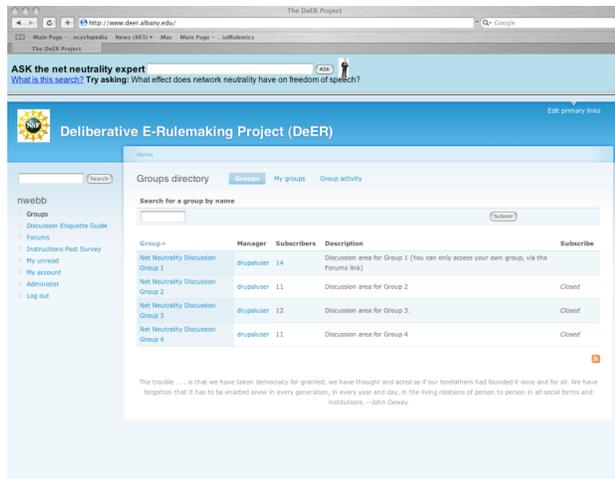


Figure 1: Screenshot of the DeER system

We mined the World Wide Web for data applicable to our first deliberation topic – Network Neutrality. We initially mined around 3.4Gb of data, which was then reduced, by removing duplicate material and HTML tags, to leave around 500Mb of final data. This data was further processed using a pipeline of NLP applications – tokenization, part-of-speech tagging, parsing, and Named Entity extraction. Finally, this data is “framed” – a vital process in HITIQA that enables us to capture some of the semantic meaning underlying passages in text. Framing is a process of relating entities in a paragraph to each other using course-grained semantic categories.

4.Pilot Findings

The pilot first phase focused on the viability of our software, procedures, and measures. One important result was that while our participants were intrigued by the QA system when it was explained to them in their focus groups, some participants defaulted to using habitual tools such as Google to learn about the topic. To circumvent this issue in the future, we will provide more detailed information on what the QA system is and its virtues, consider ways to integrate it with other useful functionality and into the discussion board, and provide a topic FAQ to help participants formulate questions. Other changes will

include more specific task instructions, larger discussion groups, seeding discussion topics, and an improved board module with a video demo. Survey measures generally fit expected factor patterns. They did well at predicting policy attitudes and knowledge, clarifying sources of change. Despite the small pilot, we found significant improvement in the sophistication of views of good government (bootstrapped $p=.03$), particularly an increased preference for flexibility in leaders and for government by checks and balances rather than centralized control.

5.Future Work

A key aim of the first phase is to acquire significant textual deliberation data in which participants outline key arguments, make counterpoints, and give examples – discovered through interactive searches at the DeER deliberation site. The next phase of processing is to analyze this data, using proven Dialogue Act Analysis techniques to automatically annotate the data with dialogue acts – labels that represent the meaning of the current utterance in context, such as a statement, question or clarification. We intend to use such DA markings to provide feedback on the quality of discussion as determined by a tested deliberative quality content coding scheme. This scheme reliably indicates the amount of agreement, disagreement, elaboration, and so forth present in a conversation. The DA analysis would also be used to identify potential paths of discussion to be presented as exemplars for other discussion groups.

6.ACKNOWLEDGMENTS

This research is based upon work supported by the National Science Foundation under Grant No. IIS-0713143.

7.REFERENCES

- [1] P. Muhlberger, "Strategic and Communicative Rationality in a Deliberative Field Experiment," presented at Midwest Political Science Association Annual National Conference, Chicago, IL, 2007.
- [2] P. Muhlberger and L. M. Weber, "Lessons from the Virtual Agora Project: The Effects of Agency, Identity, Information, and Deliberation on Political Knowledge," *Journal of Public Deliberation* (available at <http://services.bepress.com/jpd/>), vol. 2, pp. 1-39, 2006.
- [3] M. Pivato, "Pyramidal Democracy (available at: http://mpira.uni-muenchen.de/3965/1/MPRA_paper_3965.pdf)," Department of Mathematics, Trent University, Peterborough 2007.
- [4] S. Small and T. Strzalkowski, "HITIQA: High Quality Intelligence through Interactive QA (Forthcoming)," *Journal of Natural Language Engineering: Special Issue on Interactive Question Answering*, 2008.
- [5] S. Small, T. Strzalkowski, T. Liu, N. Shimizu, and B. Yamrom, "A Data Driven Approach to Interactive QA," in *New Directions in Question Answering*, M. Maybury, Ed. Cambridge, MA: MIT press, 2004, pp. 129–140.
- [6] J. Stromer-Galley, "Measuring Deliberation's Content: A Coding Scheme," *Journal of Public Deliberation*, vol. 3, 2007.
- [7] N. Webb, M. Hepple, and Y. Wilks, "Error Analysis of Dialogue Act Classification," *Proceedings of the 8th International Conference on Text, Speech and Dialogue*, Carlsbad, Czech Republic, 2005.