

Tools for Rules: Technology Transfer and Electronic Rulemaking

Stuart W. Shulman
University of Pittsburgh
School of Information Sciences
shulman@pitt.edu

Eduard Hovy
University of Southern California
Information Sciences Institute
hovy@isi.edu

Namhee Kwon
University of Southern California
Information Sciences Institute
kwon@isi.edu

Emily Huisman
University of Pittsburgh
Qualitative Data Analysis Program
j1tp12@gmail.com

Abstract

*Large-scale, email public comment campaigns in the United States necessitate the creation of new, specially tailored power tools for language. We call these applications “Tools for Rules”. This paper introduces the Rule Writers’ Workbench, a system developed specifically for dealing with the influx of public comments. We describe the context in which these tools were developed and report on our first round of usability testing. Our major finding is that a particular tool (‘DURIAN’ or **Duplicate Removal In lArge collectionN**) performs a critical function in the era of burgeoning email campaigns. We provide a glimpse into the real world testbed provided by over 540,000 emails submitted to the Fish & Wildlife Service on the proposed listing of the polar bear as “threatened” under the Endangered Species Act. Finally, we conclude with observations about the likely political considerations as tools like DURIAN become more commonplace on the desktop of regulatory rule writers.*

1. Introduction

Many people today need to interpret, structure, and rapidly master large quantities of text. Within that vast array of potential users, we address the specific task facing several thousand regulation writers who formulate the rules required to implement federal law in the United States (Coglianese, 2004; Kerwin 2003; Stoll, Lazarsky, & Herz, 2006). Our focus is on building prototype tools for rules that will prepare the ground for technology transfer from NSF-funded research to regular government use. At the same time, we also seek to improve citizen-government communication in spite of the potentially disruptive effects of large numbers of email public comments.

Some of the most useful public comments arrive at government in-boxes interspersed among thousands or tens of thousands (occasionally hundreds of thousands) of duplicative electronic form letters (Bimber, 2003). Researchers have speculated whether online public commenting will make the process better (Brandon & Carlitz, 2002; Carlitz & Gunn, 2002; Froomkin, 2004) or worse (Benjamin, 2006; Emery & Emery, 2005). For some, the periodic emergence of a large comment campaign is an outlier among more steady and manageable comment flows (ex., Balla & Daniels, 2007; Coglianese, 2006). Regardless of the trends and implications, our tools for rules seek to make it easier for government personnel and interested citizens or interest groups to find and respond to unique and meaningful comments whenever and wherever they might occur.

During the last decade, text search engines and hierarchical topic directories have become most common as tools for locating information in large text collections. Text search engines assume that users can state their information needs more-or-less accurately. Hierarchical topic directories don’t require an explicit information need; however, they are most useful when the user has a goal in mind, and undirected browsing in a large text collection is rarely productive.

What if, however, the user of a search engine or directory does not know precisely what she or he is looking for? What if that depends on the collection’s contents, or if the purpose of the collection itself is to identify both expected and unexpected elements? How then can the user effectively explore the collection from different perspectives, inspect promising items or groupings, and dispense with less useful text, to

form a reasonably detailed model of what is available and relevant? Simple search and browsing tools fall short under such conditions.

In this paper, we first provide some background information about the applied context of this research. Next, we introduce tools that were developed based on five years of input from over 200 personnel at a dozen federal agencies. Then, we report on our exploratory usability findings from a September 2006 tool test conducted in Washington, DC. We also discuss the incorporation of user feedback into recent modifications of our most successful tools along with plans for future research. Next, we report on the ongoing use of a working prototype of the Rule Writers' Workbench by a federal agency: The U.S. Fish & Wildlife Service (FWS). Finally, we conclude with some observations about the status of public comments in the context of mass email public comment campaigns.

2. Background

Computer scientists in the eRulemaking Research Group focus on text clustering, text searching, near-duplicate detection, opinion identification, stakeholder characterization, and extractive summarization (Arguello & Callan, 2007; Kwon, Shulman & Hovy, 2006; Kwon, Hovy, Zhou, & Shulman, 2006; Yang & Callan, 2005; Yang, Callan, & Shulman, 2006; Yang & Callan, 2006). Social scientists in the group are studying the impact of such tools and the Internet more generally on the process of rulemaking (Schlosberg et al., 2006; Shulman, 2004, 2005). Over the last five years, our group has collected 16 public comment datasets comprising in excess of 1,000,000 public comments on federal regulatory actions. Since our last annual report (Shulman et al., 2006), significant and regular electronic mass comment campaigns have continued.¹

In the "notice and comment" procedure created by the Administrative Procedure Act of 1946, federal agencies must invite and then process comments from the public on proposed regulations and certain other decisions. Each year several proposed regulations attract tens or hundreds of thousands of form letters and modified form letters by email (de Figueiredo, 2006; Shulman, 2003). In addition, on important rules, agencies will receive scientific studies and

legal arguments of a few hundred, or occasionally a few thousand, pages.

A decade ago, many of the form letters were straightforward to process. Except for the signature, sender address, and possibly a brief hand-written comment on the back of a postcard, they were exact duplicates of the original form letter. Form letters could be identified by size, shape, or a quick glance, and then easily sorted and counted.

In contrast, electronic form letters present new analytic challenges. The per-unit cost of generating comments is dropping sharply due to the use of commonplace Internet strategies, such as referral pages, listservs, and hypertext links. The political, educational, and organizational advantages of large scale public comment campaigns are well known and widely accepted by grassroots organizers. There are increasingly sophisticated and enticing efforts to steer large numbers of citizens to Web forms that allow groups to send subtle variations on a range of form letters as part of an awareness or advocacy campaign. What used to be relatively rare is now much more commonplace as citizens respond to impassioned and compelling interest group pleas to "Take Action Now" with their own stylized or standardized exercise of their click-though democratic voice.

When grassroots comment campaigns occur, a very small number of comments raising substantive issues (as defined by the agency) or using evidence-based arguments is mixed in among a much larger number of emotional rants or slightly modified versions of form letters. Whether a proposed rule attracts 1,000 or 1,000,000 comments, detailed manual analysis of all the comments is widely considered impractical; federal agencies face resource and staff constraints and often must act in accordance with statutory deadlines for completing a regulatory action. As a result, they often hire outside private contractors to read and summarize the comments. Simple heuristics, such as file size, the email or Web relayer metadata, and visual similarity, typically are used as a triage method to identify and set aside comments likely to have been generated by grassroots comment campaigns.

Current practice has some undesirable characteristics. It is expensive, because it is largely a manual process. One agency reported to our group on receiving bids in excess of \$250,000 and \$500,000 to process approximately 120,000 and 200,000 largely duplicative public comments. Manual review of mass public

¹ One example of a multi-campaign Action Alert center is the Earthjustice "Take Action!" site at: http://www.earthjustice.org/how_to_help/action/index.html.

comment campaigns is also likely to be error-prone. Occasionally people do add substantive issues, evidence-based arguments, or relevant personal experiences to form letters and regulators readily admit they are hard to find in larger collections that strain staff resources (Shulman, Thrane, & Shelley, 2005).

One particularly suspect current practice reported by personnel at different agencies is to sort emails by the rounded-off number of kilobytes in a file. One tool tester claimed that when files have the identical number of kbs “they are probably the same darn letter.” This logic is flawed, however, as demonstrated by Figure 1, which shows a demonstration sample of 15 files and their Windows OS-reported rounded file size. The files are all unique in length in increments of 100 words. One in three files in this example would be falsely identified as identical to another file using the size rule based on a kbs heuristic.

Name ▲	Size	Type
100.txt	1 KB	Text Document
200.txt	2 KB	Text Document
300.txt	2 KB	Text Document
400.txt	3 KB	Text Document
500.txt	4 KB	Text Document
600.txt	4 KB	Text Document
700.txt	5 KB	Text Document
800.txt	6 KB	Text Document
900.txt	6 KB	Text Document
1000.txt	7 KB	Text Document
1100.txt	8 KB	Text Document
1200.txt	8 KB	Text Document
1300.txt	9 KB	Text Document
1400.txt	10 KB	Text Document
1500.txt	10 KB	Text Document

Figure 1. Fifteen text files in increments of 100 words and their rounded size in kilobytes displayed using a Windows operating system.

Manual review invariably discards considerable detail about why some members of the public support or oppose a proposed regulation, or how the public would like to see it modified. This is information that would be useful to politicians, policymakers, and social scientists (Kerwin, 2003; Lubbers, 2006). Indeed, the notice and comment process is designed to elicit useful information, but such information does not always make it through the process as it currently exists (Coglianese, 2004). Federal agencies therefore need, but do not yet

have, better tools for analyzing and managing large amounts of text data.

Our research explores the use of information extraction and information retrieval to assist rule writers and analysts in managing large volumes of public comments. Information extraction techniques strip off email headers, salutations, signature lines, and advertising text. Text clustering algorithms identify exact duplicates, group together comments that are similar but not identical, and organize them hierarchically for browsing by rule-writers. Text-differencing algorithms identify where a person has edited a form letter so that a rule-Writers’ attention is drawn immediately to the unique part of an edited form letter.

3.0 The Tools

Our central aim is to develop tools for interpreting, structuring, and rapidly mastering large quantities of opinion-based text. Work began in earnest with a six-agency workshop in 2001 at the Council for Excellence in Government. Since then, we have conducted workshops, focus groups, and interviews involving over 200 government personnel, academics, and electronic advocacy specialists, as well as a focus group with nine citizen commenters who all sent an identical form letter. These sessions have significantly shaped and externally validated our research and tool development, as well as set the agenda for our continuing efforts. By the middle of 2006, our group finalized version 1.0 prototypes of the following three tools.

2.1. A Browse and Search Baseline

What we call Tool 1 offers only a very simple browsing facility and a relatively good full-text search engine. Fairly straightforward by design, it was prepared as a baseline of what we assumed (it turns out erroneously) that most rule writers already possessed. Tool 1 allows the analyst to browse quickly and easily through a collection of electronic comments the way one might flip through a stack of paper using familiar “Next Document” and “Previous Document” links. It also includes the Lemur Toolkit’s Indri search engine, which has a very effective document ranking algorithm and, optionally, a powerful query language.

2.2. Duplicate & Near Duplicate Detection

To help rule writers and analysts navigate enormous, often duplicative, piles of virtual

comments, Tool 2 (named ‘DURIAN’ or **Duplicate Removal In lArge collectioN** in other papers) detects duplicate and near-duplicate comments, grouping them hierarchically in message groups. It also breaks down sub-groups into similar clusters, automatically highlighting any changes (additions or subtractions) in the near-duplicates (Yang & Callan, 2005, 2006; Yang, Shulman, & Callan, 2006).

2.3. Sub-Topic & Opinion Analysis

Tool 3 provides the results of various types of machine processing of the comments. Its outputs include sub-topics, keywords, and opinions that appear in the body of the comments. This information is derived automatically from the comments and displayed using tables that allow the testers to drill down into the relevant chunks of text in a systematic manner (for example, searching all the opposing opinions for the sub-topic reasons given in them).

3. Usability Testing and Feedback

In September 2006, we convened a series of three usability tests in cooperation with the Bureau of Labor Statistics (BLS) in Washington, DC. Three successive groups of ten volunteer subjects each, drawn primarily from five federal agencies (USDA, DOT, EPA, FCC, and NSF) tested the three toolsets in mock analyses of actual public comments submitted to the EPA. A tool demo site was set up to allow access to 1,000 randomly selected comments from the full set of emails submitted to the EPA about two proposed rules. One sample of public comments was from the EPA’s mercury emissions rulemaking, and the other came from the Toxic Release Inventory Burden Reduction rulemaking.

Working in a BLS computer lab, the testers were given a brief introduction to each tool and then evaluated one tool at a time. The subjects were asked to use each tool to extract as many salient points as they could during the allotted time. First the testers were given a brief orientation that described the purpose of the tests and the types of tools to be tested. Next, for approximately 25 minutes, they worked with Tool 1, followed by Tool 2 for approximately 25 minutes, and Tool 3 for about 30 minutes. We captured their responses created while they were using each tool and thus were able to analyze how much more or less each tester was able to do with each tool.

After the test, each group of ten testers spent approximately one hour in a focus group providing feedback on the tools. The subjects also discussed the role of electronic form letters and public participation more generally. These focus groups were recorded and later transcribed, and analysis of these transcripts has advanced our work significantly. In this section we report on the usability test results based on an analysis of the focus group transcripts. First, we examine the views of the tool testers on the trends in public comment and the concerns and challenges they face in their work. Then we explore the specific observations about each tool. While we do not claim these findings generalize to the full population of rule writers, the observations are indicative of the kinds of responses we might expect as larger numbers of rule writers get access to these tools.

3.1 Form Letters and Other Concerns

In our post-tool test focus group discussions, participants were prompted to speak about their experiences with the tools but also about the changing landscape of public comment and agency responses to it. Consistent with past experiences involving multi-agency focus groups (Shulman, 2005; Shulman, Thrane, & Shelley, 2005), opinions varied widely regarding the merits of form letter campaigns and appropriate heuristics for reviewing comments. Different agencies have more or less experience with very large comment campaigns garnering tens or hundreds of thousands of emails. Some participants spoke of the strain of receiving even only a few thousand public comments. Collectively, the three groups articulated a coherent set of theories about why the influx of electronic comments presented certain problems and some benefits for the regulatory process.

One participant started a long set of observations by noting that “in the olden days, it was so much easier.” This rule writer posted samples of the various identified form campaigns on his office wall and, once familiar with its main issues, added those issues to the central list of concerns and got those comments “the heck out of our way.” The official did note that if a comment began with a familiar form letter but went for three pages, “then you might want to take a look at that.” Based on comments in earlier focus groups (Shulman, 2005), we know it is commonplace to look to overall comment length, or the length of unique text added to a form letter, as an indicator of at least the potential presence of a substantive comment. The

assumption seems to be that short comments are likely to be rants while longer comments might contain some reasoning. One person noted that “if you could sort them by how much text was added” there might be “some real substance there.” Another person agreed, elaborating that “quoting from a scientific journal—that is going to take up a lot of text.”

A different participant flatly stated that “what you are looking for in most cases are those 100-page comments that represent real substance, where they actually make suggestions ... and express themselves thoughtfully.” Not every participant thought it entirely worthwhile to search through the “onslaught of emails.” One person called it a “plague.” Another stated that lesser versions of form letters were “not that helpful.” A participant said that once you have read 500 of the same comment, you need not read anymore.

Most of the tool testers seemed to agree, however, with the assessment offered by one participant that finding meaningful comments in amongst form letter campaigns currently was a “needle in the haystack” manual task, but one for which they could be held accountable. One person described using a contractor to sort the electronic comments, and noted “probably the biggest problem we had with the project was that they misassigned things ... you don’t realize until much later that, oh cripes, I should have sent that comment to a toxicologist.”

A few comments came out strongly in favor of using the raw comment volume as a metric for reporting to management and balancing competing interests. Precisely how that volume is measured appears to be idiosyncratic, ranging from creating ad hoc weighted systems to measuring the number of comments per inch of stacked paper. One take was that the core of the form letters, rather than the unique member additions, “were probably the best written,” and it was the text of the form letters themselves that was worthy of being read.

3.2 Tool 1: Simple, Useful, and Popular

One surprising finding was that Tool 1, our presumed baseline, in fact constitutes a more usable and functional system than many rule writers currently have on their desktop. The testers found it “a bit helpful in flipping quickly ... file by file.” Others said it was “great” or that they “liked number one actually,” with the reason generally being a desire to see every comment, either by “blasting” through a virtual stack or via a user-initiated search. A tester noted

it was “obviously a very simple tool that coupled with the other tools really worked very well together.” This person noted that even when you get 25,000 comments, “you’re going to print the stack anyway ... [but with Tool 1] I don’t have to print it out now.” When reflecting on how Tool 1 could be made more effective, various testers called for a folder or bookmark function. One person suggested a “parking lot” for those comments that would need to be addressed in writing the preamble to the final rule.

3.3 Tool 3: Complex, Inaccurate, and Needing More Work

Tool 3 was judged by several users as showing significant promise, but it was seen as not yet sufficiently easy to use, nor was it accurate or fine-grained enough to be useful. One tester noted the third tool resulted in the “most trouble.” This user found, “it didn’t really help me understand ... what people were saying” and that “it was hard to put the search results in context, so I floundered a lot on that last one.” Another user complained of “too many small bites” without the surrounding text. Still other users found the categorization problematic. Sub-topic classifications with a single instance of a sub-topic combination were “more distractions than they were information.” Another theme in the critique of Tool 3 was that the sub-topic labels were too broad. A word like “environment” “didn’t have a lot of meaning ... when you get down to the nitty gritty ... it wasn’t going to be useful.”

When Tool 3 labeled some passages as “suggestions,” one tool tester found the content to be “just oppositions characterized a different way and so it was totally useless.” Another noted it was hard to see the usefulness of a tool that produced results in which “some things were miscategorized.” Comments about a better version of Tool 3 became the basis for rethinking the role of the user in a topic/opinion classification system and triggered a new proposal, submitted to NSF in late 2006 and subsequently funded in 2007. In particular, the group is now focusing on a re-designed version of Tool 3 that allows users to seed the system with specific technical or legal terms that are highly salient to the rule writers and then to interact dynamically with the clusters to refine them.

3.4 Tool 2: An Effective Timesaver

Most encouraging was the unanimous praise for the logic, function, and maturity of Tool 2, which makes it possible for a single person to review and set aside examples reflecting hundreds or thousands of identical form letters in a matter of minutes. Once the “dupes” are processed with the cluster counts displayed using Tool 2, the “near-dupes” are made easily accessible for browsing and the unique text added to a form letter is highlighted automatically.

Several testers agreed this tool was ready for use (“I would take it home now,” said one), but they also made useful observations about how it might be improved. One person stated Tool 2 is “almost ready to take out of the oven.” Some of the testers called it “the best” or “wonderful” for what one person described as “simply sorting the chaff from the wheat.” Another user spoke positively about the “clustering ... that is probably the hardest thing to do when you are sitting there with a docket full of 500 comments.”

For one user interested in the cost-saving aspect, “the duplicate matching tool is really big.” Another thought it would alter the way agencies looked at the prospect of using Regulations.Gov, the federal portal for eRulemaking. For this tester, the ability to “get rid of the bulk of them” was a way to “save your sanity.” One regulatory official spoke of the staffing implications this way:

... when you talk to someone whose job it is to actually do the intake of this stuff and they're in-near tears or near retirement, you know ... [laughter] ... people who aren't even involved in the issue themselves. They're just processing paper, and they hate it. And to give them a tool where they can quickly make piles ...when they start down this road, know that there's a light at the end of the tunnel and they have tools to deal with that, and they can give you a good product, it's an enormous benefit to them and then that becomes an enormous benefit to us.

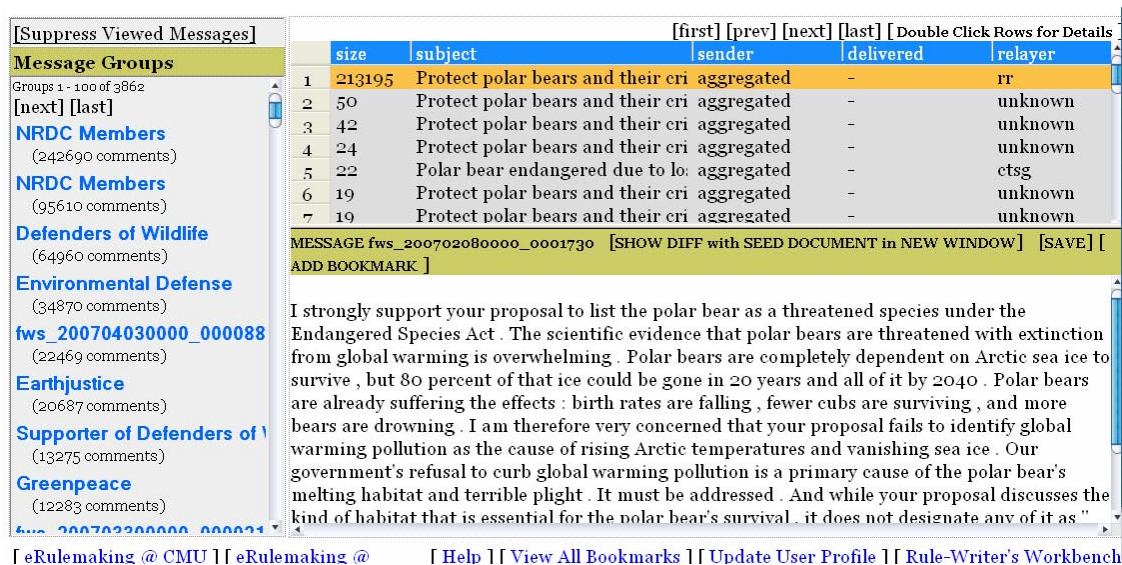


Figure 2. The DURIAN interface showing the major polar bear comment clusters.

4. A Real World Testbed: The FWS and the Polar Bear Comments

Since the time of these test sessions, the best ideas about improving Tool 2 have since been incorporated into a modified version of the Rule-Writers' Workbench. A bookmark and annotation function has been added to the interface, which allows users to return to comments they mark for more thorough review. When comments are viewed in the DURIAN

interface they are now automatically marked as read. The prototype Rule Writers' Workbench allows the registered users of the system to suppress all viewed comments, making identification of the yet to be reviewed comments easier to find. In addition, the system creates a log of the time and date that each unique comment was reviewed. This creates the possibility of a defensible legal record showing that all the comments were in fact reviewed.

Finally, the system provides accurate counts of the exact duplicates in each campaign.

At the time of this writing, the FWS is reviewing all the emailed public comments on its decision to list the polar bear as “threatened” under the Endangered Species Act. The updated Rule-Writers’ Workbench also reports the number of email addresses responsible for multiple comments. For example, we know that in the corpus of emailed public comment on the polar bear listing described below, 546,896 email comments were received between January 11 and April 9, 2007, or about 5,356 comments per day. Our system identified 261,371 email addresses that were associated with just one comment. Perhaps more interesting, 80,592 email addresses sent two comments, 34,604 addresses sent three, 6,782 sent four, and 2,607 sent five. At the long end of the tail, single email addresses responsible for 24, 71, and 163 comments respectively. Although the public comment process is not a vote, these data suggest that many of the citizens engaging in a mass public comment campaign may feel that it is a plebiscite-style form of voting and that voting more than once increases the tangible support for their cause.

A Web-based version of the Rule Writers’ Workbench, hosted by Carnegie Mellon University, currently is being used by U.S. Fish & Wildlife Service (FWS) personnel in Anchorage, Alaska, to sort and analyze the email public comments on the polar bear issue. In response to the proposed listing, many environmental groups have launched “Action Alerts” and have generated hundreds of thousands of emails. The FWS personnel are using the new prototype to review, bookmark, and annotate the full set of emails received. FWS personnel report that they spend far less time trying to find the unique comments and more time reviewing the actual passages of text inserted by interest group members. With over 80% of the comments aggregated in message groups of at least 6 exact duplicates (the largest group being comprised of 213,196 unmodified form letters generated by the Natural Resources Defense Council), one comment analyst can review every unique comment in the corpus effectively and efficiently.

5. The Future of Public Comment

New tools for rules inevitably will end up on the desktops of regulation writers. As it becomes easier to generate large numbers of comments,

tools such as these will be essential for the smooth and cost-effective functioning of the regulatory process. The APA-derived “notice and comment” process never was intended to be a rote slog through giant piles of duplicative comments sorted manually by the shape of the words on a printed page. The applied side of tools for rules research is eliminating the so-called “plague” of duplicative emails that obfuscates the role of knowledgeable public commenters shaping the views of experts in federal agencies.

Invariably, new tools for rules also will impact the way various vendors, interest groups, individual commenters, and agency personnel view the role of email in the public comment process. Efforts at manipulating the rulemaking system predate the digital communications era. It remains to be seen how more widespread use of duplicate detection and other tools under development for the Rule Writers’ Workbench will alter the tactics used by groups and vendors who created the existing mass email public comment system.

6. References

- Arguello, J. and J. Callan, “A Bootstrapping Approach for Identifying Stakeholders in Public-Comment Corpora,” Proceedings of the 8th International Conference on Digital Government Research (dg.o 2007), Philadelphia, PA.
- Balla, S.J. and B.M. Daniels, “Information Technology and Public Commenting on Agency Regulations,” *Regulation and Governance*, 1(1), 2007, pp. 46-69.
- Benjamin, S.M., “Evaluating E-Rulemaking: Public Participation and Political Institutions,” *Duke Law Review*, 55(5), 2006, pp. 893-941.
- Bimber, B., *Information and American Democracy: Technology in the Evolution of Political Power*, Cambridge University Press, New York, 2003.
- Brandon, B.H. and R.D. Carlitz, “Online Rulemaking and Other Tools for Strengthening Civic Infrastructure,” *Administrative Law Review*, 54(4), 2002, pp. 1421-1478.
- Carlitz, R.D. and R.W. Gunn, “Online Rulemaking: A Step Toward E-Governance,” *Government Information Quarterly*, 19, 2002, pp. 389-405.
- Coglianesse, C. “Citizen Participation In Rulemaking: Past, Present, And Future,” *Duke Law Review*, 55(5), 2006, pp. 943-968.
- Coglianesse, C., “E-Rulemaking: Information Technology and Regulatory Policy,” *Regulatory Policy Program Report No. RPP-05*, 2004. Available at: <http://tinyurl.com/2g2y8h>.

- de Figueiredo, J.M., "When Do Interest Groups Use Electronic Rulemaking?" White Paper prepared for the workshop eRulemaking at the Crossroads, convened at the dg.o 2006 International Conference on Digital Government Research, San Diego, CA, Proceedings available online at: <http://erulemaking.ucsur.pitt.edu/doc/Crossroads.pdf>.
- Emery, F. and A. Emery, "A Modest Proposal: Improve E-Rulemaking by Improving Comments," *Administrative and Regulatory Law News*, 31(1), 2005, pp. 8-9.
- Froomkin, A. M., "Technologies for Democracy," in Peter Shane, ed., *Democracy Online: The Prospects for Political Renewal through the Internet*, Routledge, New York, 2004, pp. 3-20.
- Kerwin, C., *Rulemaking: How Government Agencies Write Law and Make Policy*, 3rd Ed., CQ Press, Washington, DC, 2003.
- Kwon, N., S.W. Shulman, and E.H. Hovy, *Multidimensional Text Analysis for eRulemaking*, Proceedings of the 7th National Conference on Digital Government (dg.o 2006), San Diego, CA, 2006.
- Kwon, N., E.H. Hovy, L. Zhou, and S.W. Shulman, *Identifying and Classifying Subjective Claims*, Proceedings of the 8th International Conference on Digital Government Research (dg.o 2007), Philadelphia, PA, 2007.
- Lubbers, J., *A Guide to Federal Agency Rulemaking*, 4th Ed., ABA Publishing, Chicago, IL, 2006.
- Stoll, R., K. Lazarsky, and M. Herz, "Rulemaking," *Developments in Administrative Law and Regulatory Practice 2004-2005* Jeffrey S. Lubbers, Ed., ABA Publishing, Chicago, IL, 2006, pp. 135-165.
- Schlosberg, D., S. Shulman, and S. Zavestoski, "Virtual Environmental Citizenship: Web-Based Public Participation on Environmental Rulemaking in the US," in Dobson & Bell, Eds., *Environmental Citizenship: Getting From Here to There*, MIT Press, Cambridge, MA, 2006, pp. 207-236.
- Shulman, S., "An Experiment in Digital Government at the United States National Organic Program," *Agriculture and Human Values*, 20(3), 2003, pp. 253-265.
- Shulman, S., "The Internet Still Might (but Probably Won't) Change Everything," *I/S Journal*, 1(1), 2005, pp. 111-145.
- Shulman, S.W., *E-Rulemaking: Issues in Current Research and Practice*, *International Journal of Public Administration*, 28, 2005, pp. 621-641.
- Shulman, S., "Whither Deliberation? Mass Email Campaigns and U.S. Regulatory Rulemaking," *Journal of E-Government*, 3(3), 2006, pp. 41-64.
- Shulman, S.W., L.E. Thrane, and M.C. Shelley, "E-Rulemaking," in G. David Garson, ed., *Handbook of Public Information Systems*, 2nd ed., CRC Press, Boca Raton, FL, 2005, pp. 237-254.
- Shulman, S., J. Callan, E. Hovy, and S. Zavestoski, "Language Processing Technologies for Electronic Rulemaking: A Project Highlight," Proceedings of the 7th National Conference on Digital Government Research, Digital Government Research Center, May 21-24, 2006.
- Yang, H., and J. Callan, "Near Duplicate Detection for eRulemaking," Proceedings of the 6th National Conference on Digital Government Research, Digital Government Research Center, May 21-24, 2005.
- Yang, H., J. Callan, S. Shulman, "Next Steps in Near Duplicate Detection for eRulemaking," Proceedings of the Seventh National Conference on Digital Government Research, Digital Government Research Center, May 21-24, 2006.