

Revision Memo for, “Government Websites As Data: A
methodological pipeline with application to the websites
of municipalities in the United States”

September 23, 2021

Dear Profs. Copeland, Gainous, and Towner,

Thank you for your comments and for giving us the opportunity to revise and resubmit our paper. Thanks also to Reviewer 1 for their constructive comments. As we hope you will see both in the revised manuscript and in this response letter, we have taken each comment seriously and have endeavored to address each as completely as possible. Each of the comments is addressed in the order presented in the review. We should note that we are now approximately NNN words over the 5,000 word limit. We trimmed the text substantially as we revised the paper, but we think the current version provides the most effective presentation of our core contributions and the revisions suggested by the reviewers. Of course, we are happy to trim further if necessary. We do believe that addressing reviewers' comments has significantly improved the paper.

Reviewer 1

1. *My main recommendation is to add some figures/plots that show the substantive findings about the differences between Democratic and Republican mayors. You do a good job describing the findings in the text, but I think figures would also help make the findings clear to readers.*

Response: We thank the reviewer for this recommendation. We recognize that the STM results presented in Table 2 conveyed the content, direction of partisan bias, and statistical significance, but did not convey magnitude of differences/effects in any meaningful way. We have added Figure 3, in which we illustrate the magnitudes of the differences between the prevalence of topics based on the mayor's partisanship. We also added discussion of the differences illustrated in this figure. We see this as a major enhancement of our results interpretation.

2. *Also, in the conclusion can you give some more examples of other research areas and questions where this approach will be useful? I think that would increase the usage and citations.*

Response: We have added a paragraph to the conclusion in which we recap the uses of website content analysis that we reviewed in the front end of the paper, and also note and cite potential applications outside of the analysis of government websites and candidates. We give examples of politics and policy research in which researchers have analyzed (1) company websites, and (2) the websites of interest groups.

3. *Along similar lines, I hope you plan to produce step-by-step replication code and instructions about how readers can use your workflow. I think that too will increase the usage and citations of your approach.*

Response: We will make three categories of online materials available to help readers/users effectively apply the methods workflow we develop. First, we will post a conventional replication archive to Dataverse. The replication archive will include all of the code necessary to replicate the text analysis results we present, as well as the

raw text corpora that we extracted from the websites. Second, the `gov2text` package is available on GitHub. Third, we have posted an interactive tutorial on Google Colab that demonstrates the key steps in the application of our pipeline using the `gov2text` package.

4. *One minor point – there’s a typo in the abstract – “can can”*

Response: Fixed.