hagen_2019_open_data_visualizations_and_analytics_as_tools_for_policy_making

Year

2019

Author(s)

Loni Hagen and Thomas E. Keller and Xiaoyi Yerden and Luis Felipe Luna-Reyes

Title

Open data visualizations and analytics as tools for policy-making

Venue

Government Information Quarterly

Topic labeling

Fully automated

Focus

Secondary

Type of contribution

Established approach

Underlying technique

Probability-based method (LDAvis relevance threshold)

Topic labeling parameters

weight parameter λ: 1 and 0.6

Label generation

We then developed visualizations for these LDA topics using LDAvis, an open source topic modeling visualization tool

we labeled each topic based on the LDAvis visualization results.

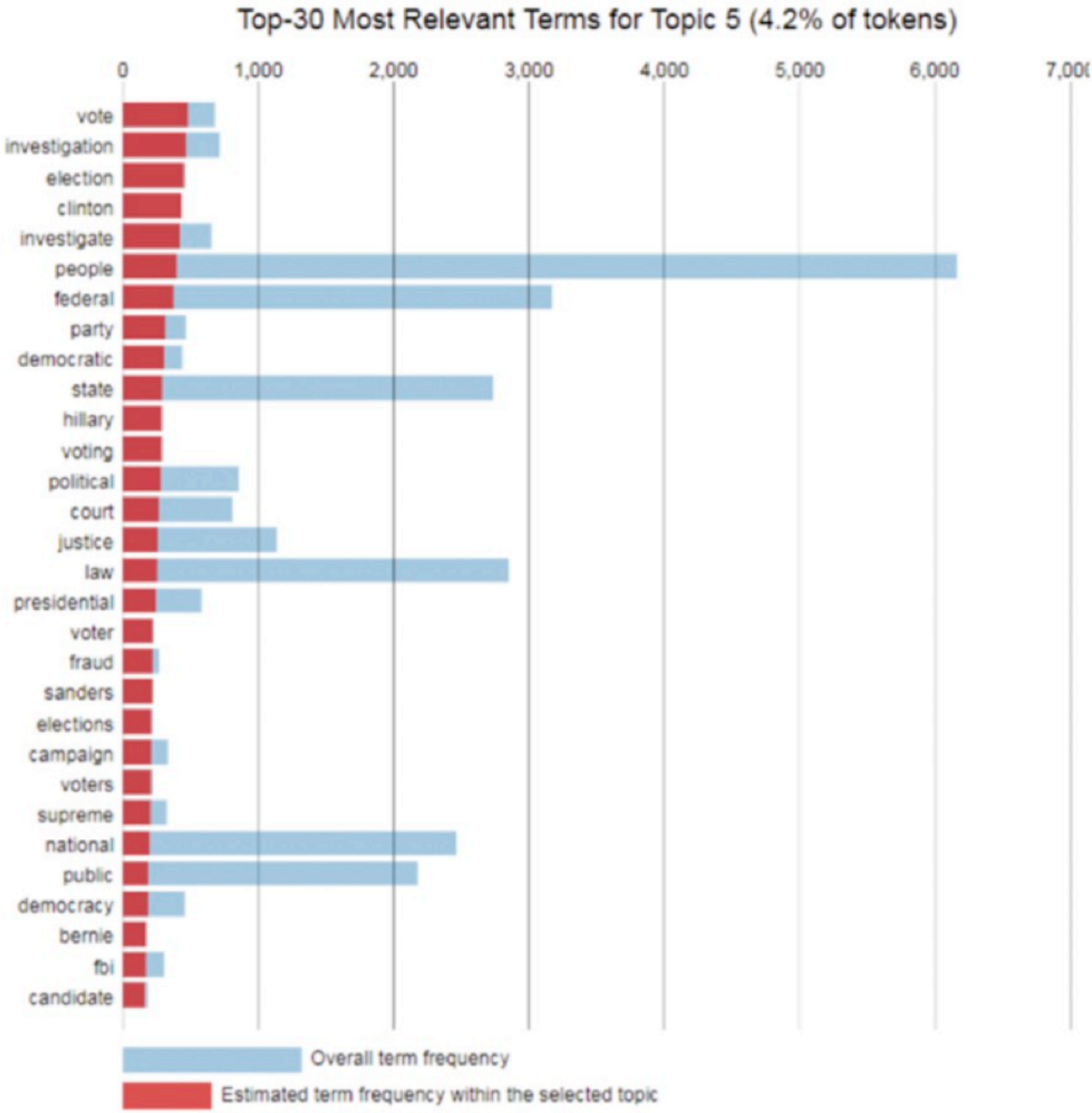
The topic words were sorted in descending order based on the estimated term frequency within the selected topic, which informs topic words that are highly relevant to the specific topic. The relevance of a term to topic is given by a weight parameter λ. Topic words displayed in Fig. 3(a) are acquired using λ = 1. Topic words displayed in Fig. 3(b) are results from using λ = 0.6,

labels are selected from the top 10 topic words (except Police & BLM) displayed by LDAvis (relevance parameter λ = 0.6)

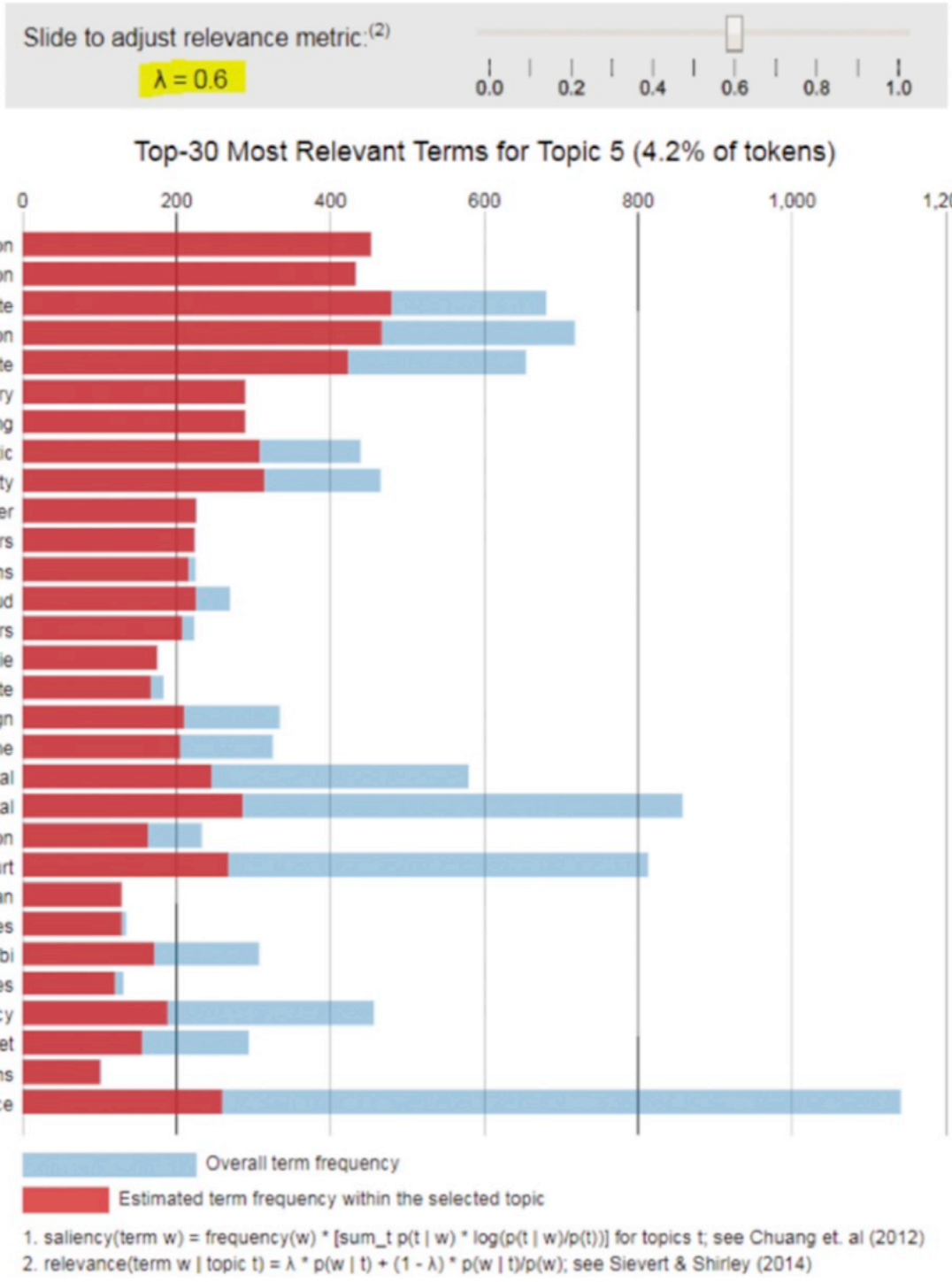
Slide to adjust relevance metric:[2]

λ = 1

0.0 0.2 0.4 0.6 0.8 1.0



1. saliency(term w) = frequency(w) * [sum_t p(t | w) * log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
2. relevance(term w | topic t) = λ * p(w | t) + (1 - λ) * p(w | t)/p(w); see Sievert & Shirley (2014)



Motivation

To help the interpretation and further analyses

Topic modeling

LDA

Topic modeling parameters

Nr of topics: 30

Nr. of topics

26 (making sense for human interpretation)

Label

Single word extracted from LDAvis

Label selection

\

Label quality evaluation

\

Assessors

\

Domain

Paper: Policy

Dataset: Policy

Problem statement

This article reports opinions found in “We the People” petition data using topic modeling and visual analytics. It provides an assessment of the usability of the visual analytics results for policy making based on interviews with data professionals and policy makers. Major contributions of this study include: (1) sug gesting viable visualization tools for analyzing textual data for policy making, and (2) suggesting how to lower barriers to adoption by increasing usability.

Corpus

Origin: WtP website

Nr. of documents: 4985

Details:

- between September 22, 2011 (the initiation date), and July 12, 2016

Document

petition title and its corresponding rationale

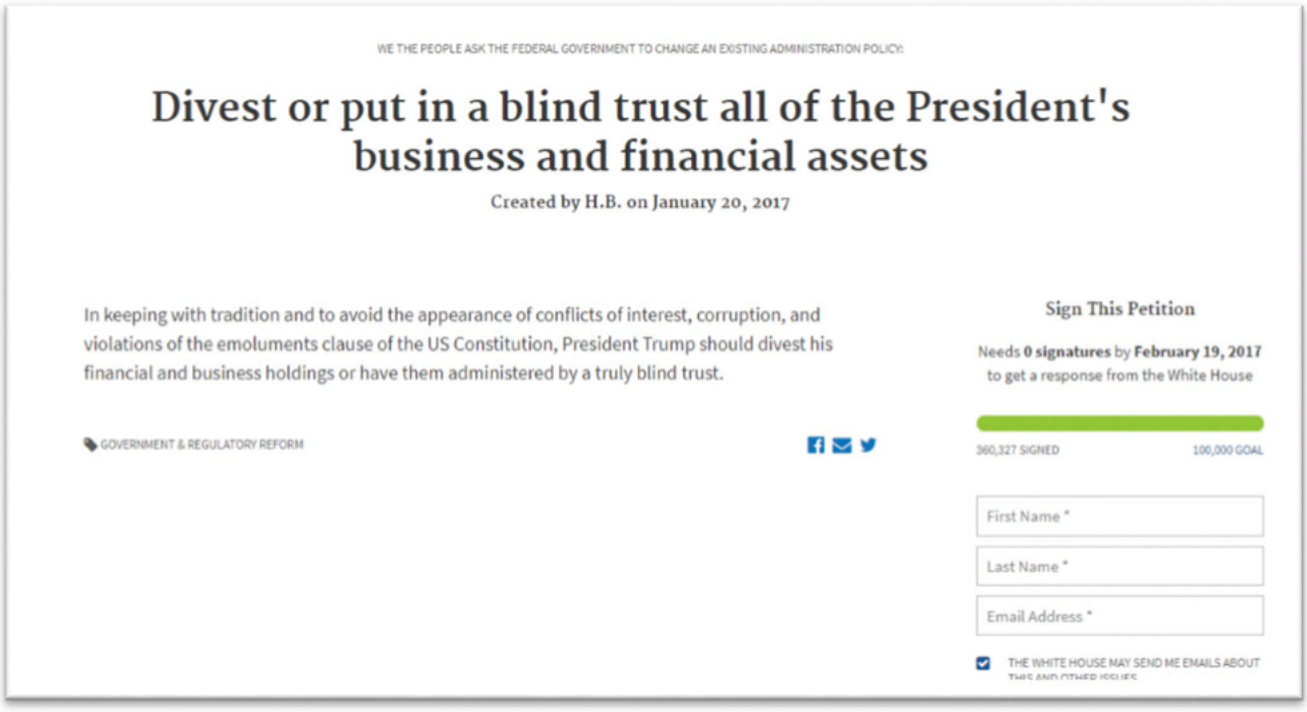


Fig. 1. An example of an WtP petition.
Note: The first two lines (bold and large font) are the title of the petition, and the rest of the text is the rationale of this petition.

Pre-processing

- converted all texts to lower case
- normalized white spaces

- eliminated punctuations
- non-al- phanumeric characters
- removed short words of only one or two characters

```
@article{hagen_2019_open_data_visualizations_and_analytics_as_tools_for_policy_making,
    abstract = {Government agencies collect large amounts of structured and unstructured data. Although these data can be used to improve services as well as policy processes, it is not always clear how to analyze the data and how to glean insights for policy making, especially when the data includes large volumes of unstructured text data. This article reports opinions found in ``We the People'' petition data using topic modeling and visual analytics. It provides an assessment of the usability of the visual analytics results for policy making based on interviews with data professionals and policy makers. We found that visual analytics have potentially positive impacts on policy making practices. Experts also articulated potential barriers regarding the adoption of visual analytics tools, and made suggestions. Potential barriers included insufficient resources in government agencies and difficulty integrating analytics with current work practices. The main suggestions involved providing training and interpretation guidelines along with the visual analytics tools. Major contributions of this study include: (1) suggesting viable visualization tools for analyzing textual data for policy making, and (2) suggesting how to lower barriers to adoption by increasing usability.},
    author = {Loni Hagen and Thomas E. Keller and Xiaoyi Yerden and Luis Felipe Luna-Reyes},
    date-added = {2023-04-20 21:25:06 +0200},
    date-modified = {2023-04-20 21:25:06 +0200},
    doi = {https://doi.org/10.1016/j.giq.2019.06.004},
    issn = {0740-624X},
    journal = {Government Information Quarterly},
    keywords = {Policy informatics, Policy analytics, Open data, Topic modeling, Visual analytics, Usability testing},
    number = {4},
    pages = {101387},
    title = {Open data visualizations and analytics as tools for policy-making},
    url = {https://www.sciencedirect.com/science/article/pii/S0740624X1830368X},
    volume = {36},
    year = {2019}}
```

#Thesis/Papers/FS