

Displaying unbiased information in the Living Voters Guide

CSE 573 Project by Allison Obourn



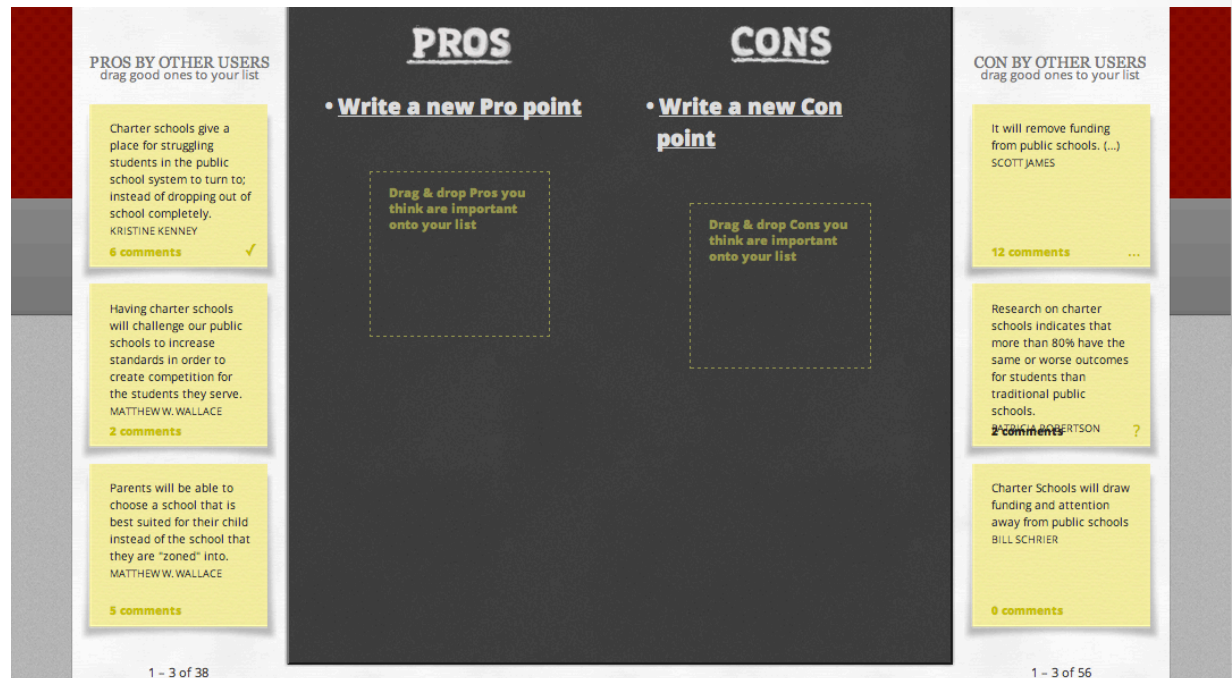
The Living Voters Guide

The screenshot shows the homepage of the Living Voters Guide. At the top right, there is a link for "Sign In / Create Account". The main header features a collage of diverse people's faces with the "Living voters guide" logo in a red circle. Below the header, a sub-header reads "A guide powered by the wisdom of Washington voters" with a "Like" button. A prompt asks users to "Click a ballot measure below. Then decide for yourself, with the help of fellow voters." There are eight red buttons representing ballot measures: 1-1240 CHARTER SCHOOLS, R-74 SAME-SEX MARRIAGE, I-502 MARIJUANA LEGALIZATION, I-1185 2/3 MAJORITY FOR TAX INCREASE, J-8223 UNIVERSITY INVESTMENTS, A-1 DEDUCTIONS FOR BANKS, J-8221 STATE DEBT, and A-2 PETROLEUM TAX EXTENSION. Below these is a search bar for "Your address" with an "Unlock" button and a link to "browse by county". A note mentions "fact-checking by The Seattle Public Library". The footer contains sections for "INFORMATION FOR VOTERS" (listing various organizations), "PAST RESULTS 2010-2011", "PARTNERS" (listing the Center for Comm & Civic Engagement and The Seattle Public Library), "FUNDERS" (listing the National Science Foundation and Google Research Award), "CO-PROMOTERS" (thanking co-promoters), and "POWERED BY CONSIDEREE" (encouraging organizational use).

- Provides information about ballot measures
- Promotes deliberation

Pro and Con Lists

- Site suggests pros and cons to users
- The pros and cons displayed effect what the user puts in their list



- Selection must be unbiased

Goals

- High level:
 - Suggest unbiased groups of points
 - Suggest points covering the topic space
- Specific:
 - Create a summary of the points in the system
 - Cluster the points and pick a point from each cluster

The intent of this initiative is good; but the specific requirements are rigid and unrealistic on the given time scale and current budget.

To comply with 1163, the state would need to hire three additional auditors and five additional fraud investigators that it can't afford.

My project

- Create similarity matrices based on different parts of the text
- Evaluate which works best

2.220446049250	0.177667721355	0.17278314860	0.163033491353
0.123060320609	0.2366009557440	0.188274539339	0.262898047091
0.131345771030	0.2693995340550	0.254049644135	0.140338427424
0.332347629498	0.2827661665480	0.1969316628130	0.251764726721
0.177667721355	-2.22044604925	0.164801292224	0.14693501793
0.137118495245	0.2374770003140	0.194451821588	0.296954501463
0.154140039915	0.247182626573	0.315748057909	0.12208753828

Method

- Bag of words approach
 - Tf-idf: term frequency – inverse document frequency
- Inputs
 - A short list of stopwords removed
 - Synonyms added and a short list of stopwords removed
 - Only nouns and verbs considered in similarity metric
 - Synonyms added and only nouns and verbs considered in similarity metric

Results

- Which text input causes Tf-idf to produce the similarity matrix closest to what a person would produce?
- Method: comparing resulting matrices with matrix generated by people voting for point similarities
- Visualize results by using the matrices to cluster the points
 - The generated clusters are very different
 - All of the results are bad!