How A Bill Became A Law

Visualizing Text Reuse in US Legislation: Patient Protection and the Affordable Care Act (PPACA) and all other 111th Congress Bills

Problem

Only a small portion of the bills that are introduced in US Congress become law. However, bills may contain policy ideas from past bills. In an attempt to trace these policy ideas from bill to bill, political scientists are using alignment methods from gene sequencing, the Smith-Waterman Algorithm, to identify sections of different bills with similar text. We seek to create an interactive tool that allows users to explore text alignments between sections of one bill, the PPACA (Obamacare), and all other bills introduced during the 111th Congress.

Approach

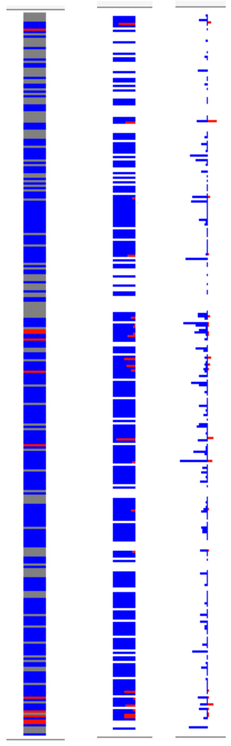
We used a left-to-right, three layer approach to allow users to see an overview of all alignments in the PPACA and direct them to interesting sections to explore.

Motivation

Visualizing a large number of text comparisons is a difficult task. A straightforward adjacency matrix is too large and sparse to be tenable, and too bipartite to show as a graph. The Political Science researchers supplying our data have made a static visualization but we felt it lacked some essential information. The graph strips out all unmatched sections of the PPACA, unintentionally implying there were no unmatched sections. The graph does not show the actual matched text itself. Lastly, a harsh threshold had to be applied to make the graph readable, meaning many potentially viable matches are not shown. We felt we could make a visualization that communicated the same ideas as the graph while also addressing the problems listed above.

One Screen Overview

The first layer of our visualization shows an overview of all alignments in the 400+ sections of the PPACA while still fitting on the screen without scrolling. We went through several different approaches for what to show within each section, including an aggregated "Democrat" or "Republican" rating for each section based on the political association of it's oldest matching text and displaying the ratio of republican to democratic text matches. Our current implementation is two bar charts showing the number of matching democratic and republican texts. We felt it was a more honest representation than aggregation and displayed more information than a ratio.

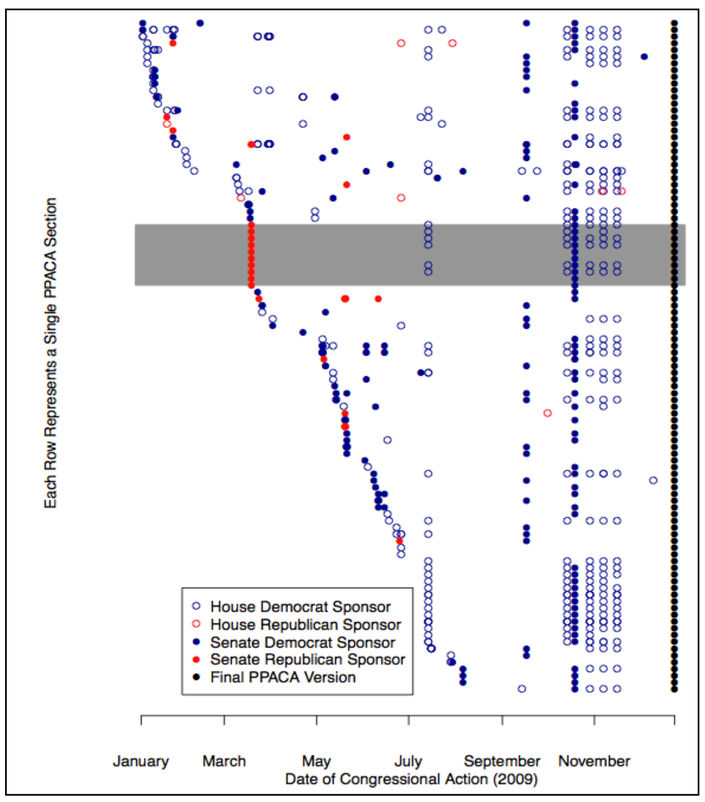


Alignment Chart

Once a user has selected a specific section of the PPACA, they see a graph of all alignments within that section. Alignment location within the text is shown on the vertical axis and matches are sorted by introduction date of the source text. We found many of the alignments came from the same bill or bills introduced at the same time. Thus, we could not show alignments on a linear time scale and had to settle for chronological sorting. To better communicate when bills were written we group bills into 3-month periods.

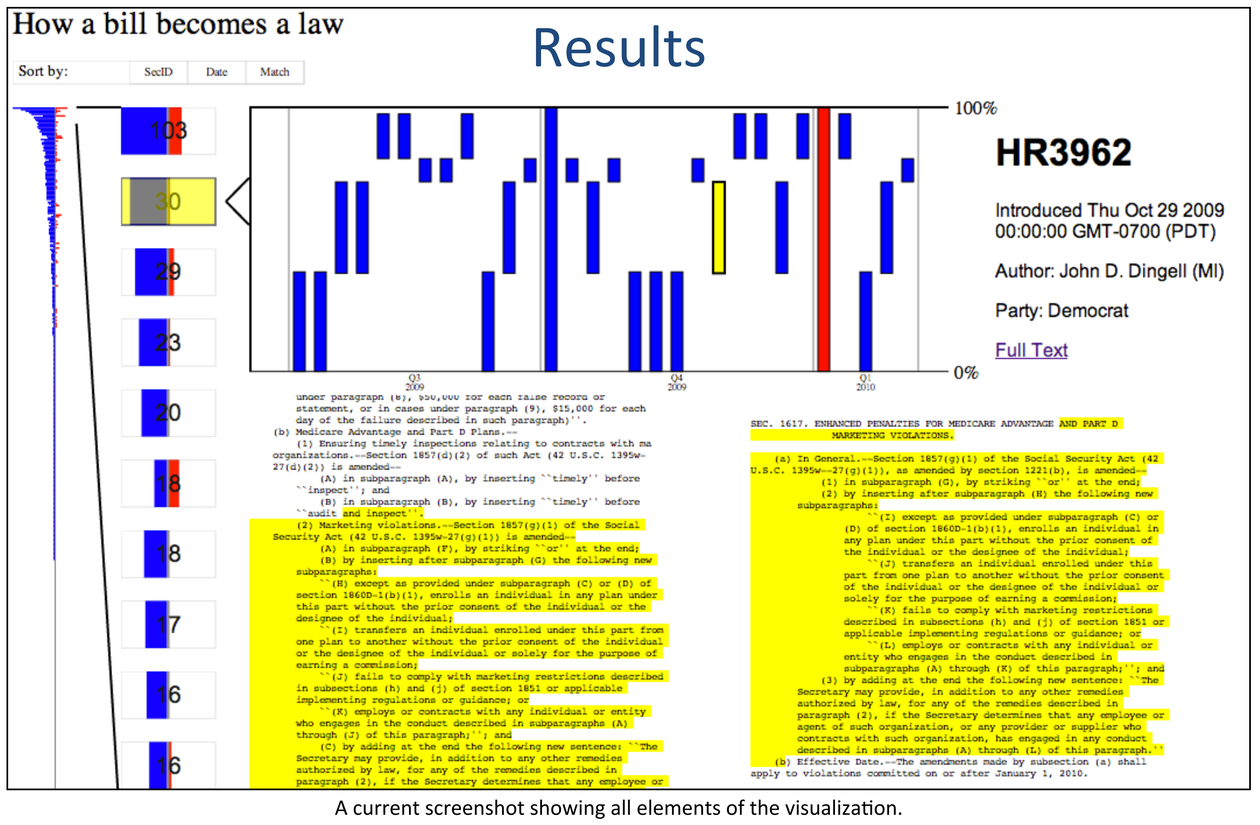
Text Alignment

When a user selects a specific alignment from the alignment chart, they see both full section texts with the matching text highlighted. Alignment highlighting allows users to easily identify aligned text and judge whether or not policy ideas are being reused between two bills. Many of the alignments are simply boilerplate text which contain no actual policy. Political Science researchers are still attempting to identify and exclude these boilerplate sections with machine learning.



Previous visualization work by John Wilkerson and Nick Stramp

Results



A current screenshot showing all elements of the visualization.

Future Work

We plan to gather feedback and work on styling for the rest of the project period. We also intend to make the software useable for any bill with the same data format, making a useful long-term tool for our Political Science colleagues.

Acknowledgments

Many thanks to John Wilkerson and Nick Strampt for sharing their data with us and answering all of our questions. Thanks to Jeff Heer and Kanit Wongsuphasawat for teaching and TAing an excellent class

Contact Info

Rob Thompson robthomp@uw.edu
Lucy Williams lgw2@uw.edu
Sam Wilson 1samw1@gmail.com

Based on work by John Wilkerson and Nick Stramp:

John Wilkerson, David Smith and Nick Stramp. 2014. "Tracing the Flow of Policy Ideas in Legislatures: A Text Reuse Approach."

QR code for our visualization