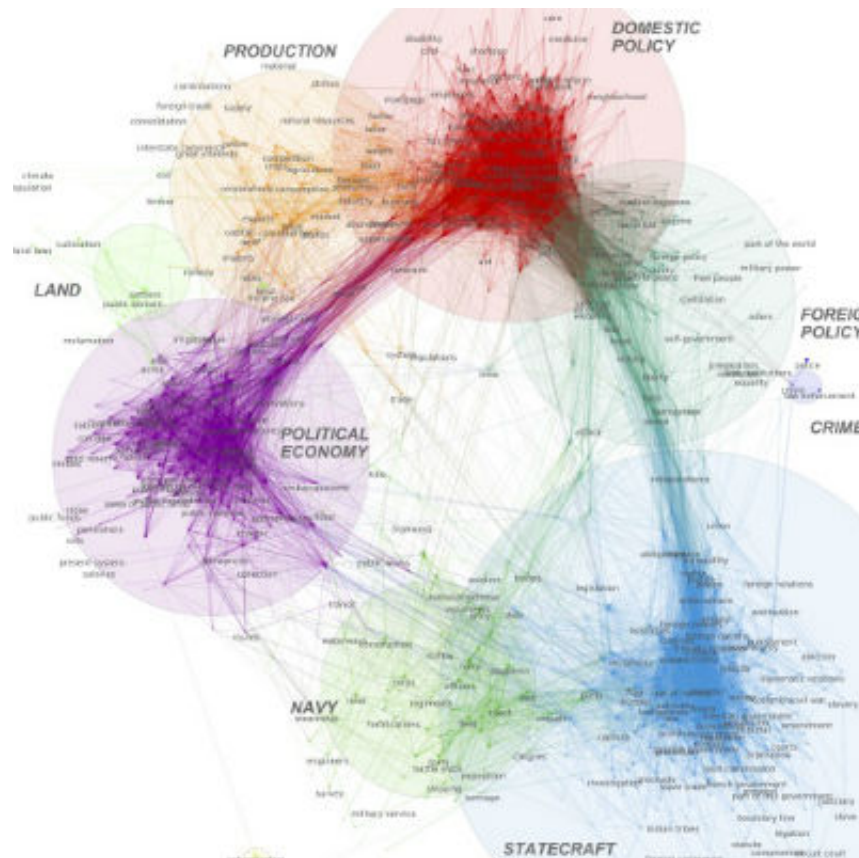


# Big data analysis of state of the union remarks changes view of American History

August 10 2015



Researchers used computational techniques to map recurring words and their relation to each other across 224-years of State of the Union remarks. Viewed as a network, the words point to common themes and disruptions in political discourse. Credit: Courtesy of the authors

No historical record may capture the nation's changing political

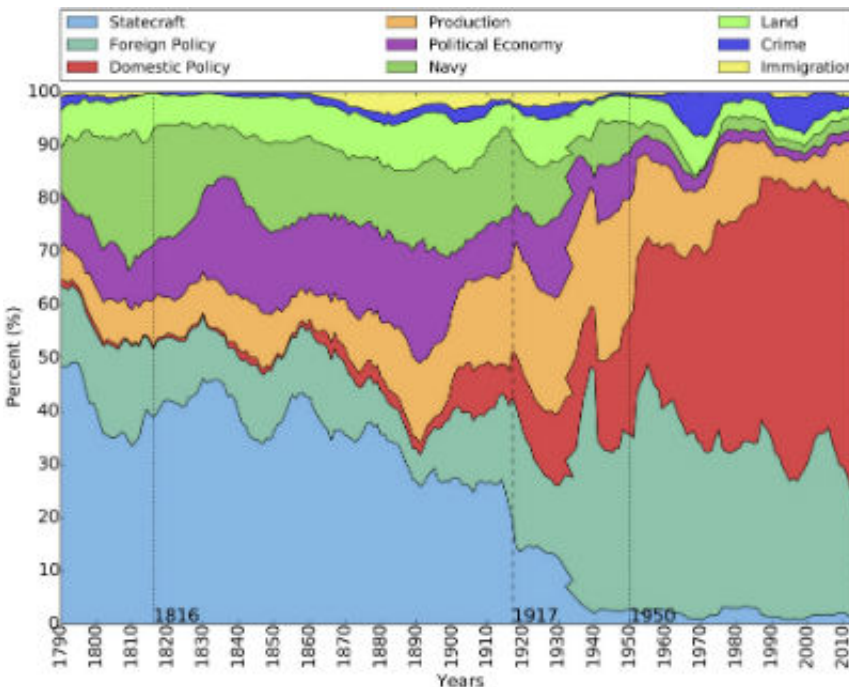
consciousness better than the president's State of the Union address, delivered each year except one since 1790.

Now, a computer analysis of this unique archive puts the start of the modern era at America's entry into World War I, challenging histories placing it after Reconstruction, the New Deal or World War II. A team of researchers at Columbia University and University of Paris published their results this week in the *Proceedings of the National Academy of Sciences*.

Though discussion of industry, finance and [foreign policy](#) dominate the record year after year, the study shows that modern political thought, defined by nation building, the regulation of business and the financing of public infrastructure, emerges with a sharp line after WWI.

"We know what constitutes modern political thinking but until now have been unable to say exactly when it originated," said the study's senior author, Peter Bearman, a sociology professor at Columbia and a member of the Data Science Institute. "Overall, our study finds striking continuity throughout the State of the Union address and a few major changes. Surprisingly, we find that key moments of disruption were unrelated to changes in the mode of delivery."

The researchers developed algorithms to analyze the nearly 1.8 million words used by American presidents in their State of the Union addresses, from George Washington's penned remarks in 1790 to Barack Obama's televised speech in 2014. By identifying how often words appeared jointly, and mapping their relation to other clusters of words, researchers were able to infer the dominant social and political discourses of the day and chart their evolution over time.



The researchers place the shift to modern political discourse in 1917, as keywords focused on the economy, public spending, government regulation and nation building emerge (covered by "Domestic Policy" in red and "Foreign Policy" in dark green). Credit: Courtesy of the authors

They were surprised to see 1917 jump out so clearly. As the United States joined Allied forces in the war against Germany, the researchers found a new set of terms recurring in the State of the Union address. On the topic of foreign policy, "democracy," "unity," "peace" and "terror" emerged as keywords, replacing older notions of statecraft and diplomacy. By the 1940s, a cluster of terms centered on the Navy, perhaps signifying an isolationist foreign policy, all but disappears. "Suddenly the U.S. is no longer an island," said Bearman.

The researchers also found a shift in terminology around domestic policy, as a new conversation over the size of government and its role in regulating the economy and providing equal opportunity emerges.

Though the underlying focus stays the same, keywords such as the "Treasury," "amount" and "expenditures" are replaced by "tax relief," "incentives" and "welfare" as America transitions from a classical political economy to the modern welfare state.

"Though the language and entire discourse of governance changes, the conversation streams remain continuous," said study lead author Alix Rule, a graduate student at Columbia.

One challenge in studying two hundreds years of political discourse is that language naturally evolves. Words may stay the same but acquire new meaning; new words may come into use to describe recurring themes. The study uses network-analysis techniques developed by coauthor Jean-Philippe Cointet, a physicist at the University of Paris, to highlight the meaningful changes, and show how some political topics morph into similar topics with common threads while others peter out and die.

The techniques allowed the researchers to capture the meaning of words in relation to other words and in the broader context of evolving topics. The study found that in America's early history the word Constitution was most commonly associated with "people." After the Civil War, "Constitution" became most closely linked to "state," soon to be linked to "law" during WWI and WWII, before becoming associated with the word "people" again in the 1970s. What the word Constitution means at any given time, the authors argue, depends on the words it is linked to.

David Blei, a statistician and computer scientist at Columbia's Data Science Institute who was not involved in the research, says the study pushes the boundaries for statistical machine learning of language. "The authors have developed an impressive and ambitious methodology for revealing the flow of thought and sentiment within a sequence of political texts," he said.

**More information:** Lexical shifts, substantive changes, and continuity in State of the Union discourse, 1790–2014,  
[www.pnas.org/cgi/doi/10.1073/pnas.1512221112](http://www.pnas.org/cgi/doi/10.1073/pnas.1512221112)

Provided by Columbia University

Citation: Big data analysis of state of the union remarks changes view of American History (2015, August 10) retrieved 24 April 2024 from <https://phys.org/news/2015-08-big-analysis-state-union-remarks.html>

<p>This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.</p>
--