Energy Policy in Flux: An NLP Approach to Media Discourse Before and After the Russia-Ukraine Conflict

Priscila Stisman

April 27, 2025

Index

Introduction

- **Hypothesis:** The onset of the Russia-Ukraine war acted as a catalyst for increased attention on energy policy, driven by rising concerns over natural gas supply, oil prices, and energy security
- **Research Question:** To what extent, and in what ways, has the Russia-Ukraine war shaped public attention and perspectives on energy policy and security?
- **Data:** 18,293 energy-related New York Times news articles from 2021 to 2023

Corpus

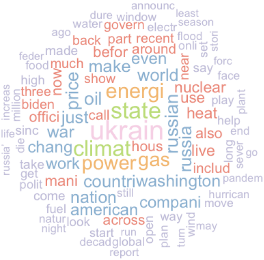
- Source: New York Times article metadata (2000–2025)
- Time period: 2021 to 2023 (126k+ articles)
- Filtered for energy-related articles (**18,293 articles**)
- Key variables: publication date, headline, abstract, and lead paragraph

Methods

- **Dictionary-Based Analysis:** Evaluate shifts in sentiment of energy policy discussions before and after the war
- **LDA Topic Modeling**: Identify how energy policy discussions evolved in response to the Russia-Ukraine war
- **Interrupted Time Series Regression**: Assess the impact of the war on the energy policy coverage in news + identify structural breaks in trend
- **Structural Topic Modeling**: Uncover key themes in energy-related news coverage and analyze how their prevalence has evolved over time

Exploratory Data Analysis

2021



2022

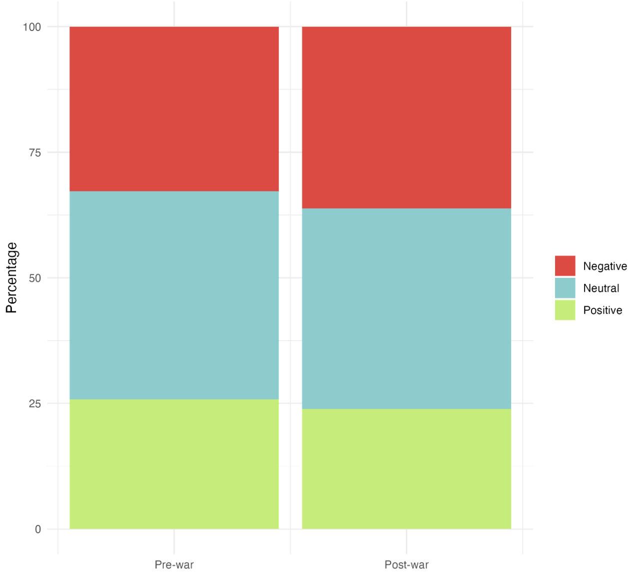
2023

Year	Top 10 Words
2021	climat, state, chang, washington, biden, vaccin, work, pandem, american, make
2022	ukrain, state, climat, energi, power, gas, russian, oil, price, russia
2023	state, climat, wordl, power, make, world, nation, even, work, window

Sentiment Analysis

- Corpus subset: **Energy Policy** (6,664 articles)
- Analysis of NYT article **headlines** before and after the war
- VADER

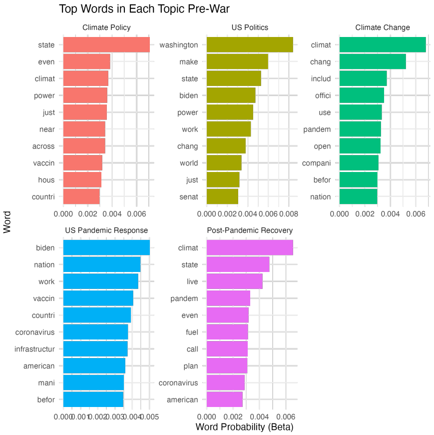
Shifts in Sentiment on Energy Policy

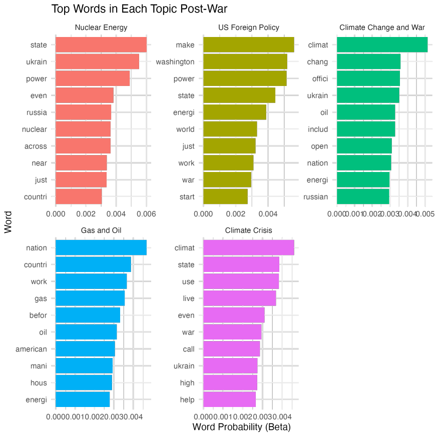


Latent Dirichlet Allocation

- Corpus: all energy-related articles (18k+)
- Explore topics in **lead paragraph** by year and before and after the war
- Used coherence and perplexity scores to calculate optimal K (K* = 5, local minimum) (see appendix).

Topic Modeling: Before vs. After the War





Interrupted Time Series

Linear regression model to estimate the effect of the war on the frequency of energy-related discussions in the news

- Is there a significant change in Energy Score following the war?
- Are there underlying time trends independent of the war? Is there any structural break?

$$EnergyScore = \beta_0 + \beta_1 War \tag{1}$$

$$EnergyScore = \beta_0 + \beta_1 War + \beta_2 Time + \beta_3 War.Time \tag{2}$$

where:

$$EnergyScore = \frac{\text{energy keywords in headline} + \text{abstract} + \text{lead paragraph}}{\text{total word count}}$$

Interrupted Time Series Results

Model 1: No Time Trend

Variable	Energy Score
Intercept	3.207***
	(0.039)
War	-0.0533
	(0.0535)
F(1,9295) = 0.9937	
p-value = 0.3189	

Model 2: With Time Trend

Variable	Energy Score
Intercept	3.334***
	(0.0832)
War	0.0009
	(0.1052)
Time Trend	0.0006
	(0.0003)
War*TimeTrend	-0.0014***
	(0.0004)
F(3,9295) = 5.162	
p-value = 0.0015	

Chow Test	Statistic	P-value
	6.005	0.0025
Reject H_0 of no structural break		

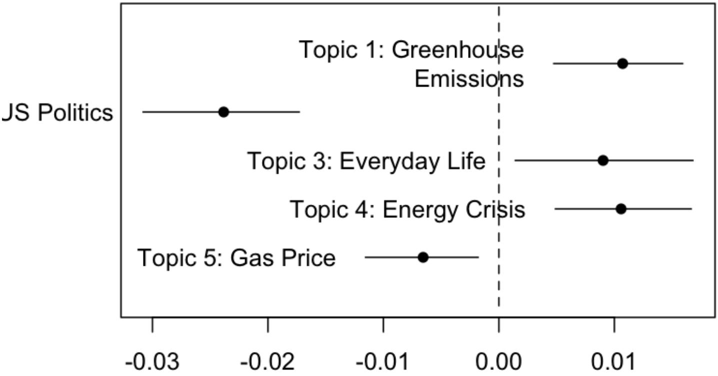
Structural Topic Models

- **Uncover latent topics within energy-related news** and estimate their temporal evolution and variations in response to the war
- **Evaluate the impact of time and the war on the prevalence** of specific themes, as integrates covariates.
- Use Spline Regression to allow for nonlinearities
- Use K*=5

STM: Topic Results

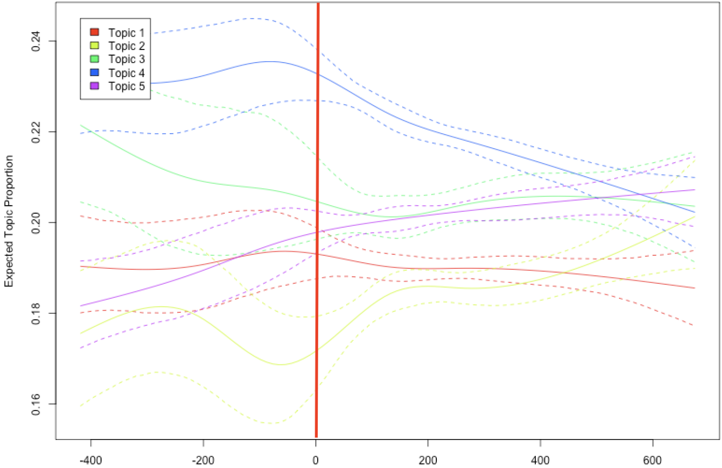
Topic	Highest Probability Words	FREX Words
1: Greenhouse Emissions	new, energi, like, week, countri, first, even	carbon, energi, inflat, look, start, forc, russian
2: US Politics	climat, presid, unit, washington, nuclear, nation, biden	biden', republican, presid, democrat, tax, washington, polici
3: Everyday Life	time, home, can, wednesday, hurrican, tuesday, govern	room, grill, recip, meat, ice, ingredi, mu-sic
4: Energy Crisis	year, one, chang, oil, day, ukrain, will	oil, fuel, will, day, world', crisi, one
5: Gas Price	state, said, gas, peopl, price, thursday, month	price, peopl, electr, rain, high, area, thursday

War Effect on Topic Prevalence



Time Effect on Topic Prevalence

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
Greenhouse Emissions	US Politics	Everyday Life	Energy Crisis	Gas Price



Conclusion

- Following the onset of the war, energy-related topics became significantly more prominent. However, their salience appears to decline gradually over time, suggesting a normalization or shift in media focus.
- Next steps include incorporating word embeddings to capture the nuanced meanings and contextual usage of energy-related terms, providing deeper insight into evolving discourse.