Putin's Talks
Winter 2026

Cyprien Fourcroy
WUT
01202818@pw.edu.pl

Katarzyna Leniec
WUT
01151680@pw.edu.pl

Jan Pogłód
WUT
01171172@pw.edu.pl

Michał Puścian
WUT
01151677@pw.edu.pl

Supervisor: Anna Wróblewska
Warsaw University of Technology
anna.wroblewska1@pw.edu.pl

Abstract

This project utilizes Natural Language Processing (NLP) methods to analyze the political speeches of Vladimir Putin, focusing on the automatic detection of topics, emotions, propaganda rhetoric, and temporal shifts in discourse. The investigation is structured around six research categories, ranging from descriptive statistics and contextual analysis to critical interpretations and robustness tests. In addition to reviewing relevant literature and open-source datasets, this study evaluates the efficacy of state-of-the-art Large Language Models (LLMs) and traditional NLP techniques in the domain of political discourse analysis.

1 Introduction

In this project, we want to use Natural Language Processing methods to analyze political speeches, especially those of Vladimir Putin. NLP gives us tools to automatically study topics, emotions, propaganda rhetoric, and how rhetoric changes over time. In this part of the report, we review the literature, available models and tools, and open-source datasets.

2 Proposed Analysis Questions

Our investigation is structured around a set of research questions grouped into six categories. The STATISTICS category includes descriptive measures of the speeches (counts, proportions, etc.). The CONTEXT category addresses external factors such as events, audience, and setting. The CHANGE OVER TIME category considers temporal trends in the discourse. The COMPARISON category involves contrasting Putin's speeches with other speeches or contexts. The INTERPRETATIONS category focuses on the substantive meaning and implications of observed patterns. Finally, CRITICAL TESTS involve validation checks and robustness analyses. Below all proposed analysis questions under these categories and proposition of methods and models to answer them are listed.

STATISTICS

- Which three economy-related terms appear most often? *(Word frequency analysis, economy lexicon)*
- How many times does the word "modernization" appear in a military context? *(Keywords co-occurrence, custom lexicons)*
- How often does he speak about "friendship" in relation to China? *(Keywords co-occurrence, custom lexicons)*
- In which years do the most references to World War II occur? *(Word frequency analysis, custom lexicon)*

CONTEXT

- What metaphors does Putin use toward the USA, Ukraine, Poland, and Germany? *(Transformer-based metaphor model / embedding similarity)*
- In references to Poland, are historical or contemporary contexts more frequent? *(Topic modeling, BERTopic)*
- In what context does "Poland" most often appear? (enemy, partner, ally, neighbor) *(Zero-shot classification)*

CHANGE OVER TIME

- How does the image of Ukraine (Poland, Germany, USA) change in his speeches from 2000 to 2024? *(Sentiment analysis, clustering, RoBERTa)*

- When do numerous references to the "Russian world" (russkiy mir) begin to appear? *(Keyword frequency over time + HuggingFace NER)*
- When do themes of "sovereignty" begin to dominate foreign policy discourse? *(Topic modeling over time, BERTopic)*
- In which years do economic topics appear most frequently? *(Word frequency analysis, economy lexicon)*

COMPARISON

- How does the narrative toward Poland differ from that toward Germany? *(Sentiment analysis, RoBERTa)*
- Does he speak more positively about China than about India? *(Sentiment analysis, RoBERTa)*
- How does the tone toward the USA compare with that toward the European Union? *(Sentiment analysis + topic modeling, RoBERTa / BERTopic)*
- How does the image of Germany differ between 2003 (Iraq War) and 2014 (Ukraine crisis)? *(Sentiment analysis, RoBERTa / BERTopic)*

INTERPRETATIONS

- What are Putin's most common arguments for strengthening the army? *(Topic modeling, BERTopic)*
- How does he construct the image of the "enemy"? *Sentence-BERT + Clustering (DBSCAN)*
- What historical events does he use to legitimize actions toward Ukraine? *(Topic modeling, BERTopic)*
- What elements of the "Great Russia" myth recur in his speeches? *(Regex search + NER / frequency counts)*
- How does he portray Russia's role in the world – as a defensive or expansionist power? *(Sentiment analysis + topic context, RoBERTa / BERTopic)*

CRITICAL TESTS

- Summarize the speech from date X *(LLMs)*
- Has Putin ever spoken about event Z? *(Keyword search + sentence extraction / LLMs (Refusal Rate))*
- Provide quotes where he describes Poland in historical terms. *(Keyword search + sentence extraction, spaCy / Regex)*
- What three different arguments does he invoke when speaking about sanctions? *(Topic modeling + co-occurrence, BERTopic / embedding clustering)*
- List the passages in which he refers to Lenin or the USSR. *(Keyword search + sentence extraction, Regex / spaCy)*

3 Literature Review

3.1 Literature and Research on NLP in Political Discourse

Analysis of Putin's rhetoric. In the 2024 paper "Analyzing Russia's propaganda tactics on Twitter using mixed methods network analysis and natural language processing: a case study of the 2022 invasion of Ukraine", Alieva et al. [2] investigate Russian propaganda discourse on Twitter during the 2022 invasion of Ukraine. The authors propose a comprehensive analytical pipeline to identify dominant topics, influential actors, and the most impactful messages contributing to the spread of disinformation narratives. Their approach combines network analysis, natural language processing techniques, and qualitative analysis.

Topic modeling. In the report "Unpacking Russian Presidential Speech Patterns with Machine Learning" [1], the authors apply Latent Dirichlet Allocation (LDA) to Vladimir Putin's speeches in order to identify dominant topics, including energy and international relations.

Propaganda detection. In a paper by Martino et al. (2020) [8], the authors describe a shared task focused on detecting propaganda in news articles. The competition received 44 submissions, and the paper provides a detailed analysis of the architectures, methods, and results obtained by participating systems. These findings offer valuable guidance by highlighting common challenges and effective approaches, and they serve as a useful starting point for model development. More broadly, online propaganda has become an increasingly important topic in NLP research. For example, the HQP dataset contains manually annotated propaganda texts and can be used to train models for propaganda classification, as discussed in "Large Language Models for Propaganda Detection".

3.2 Tools and Open Source Methods

To analyze Putin's speeches, we plan to use these tools and pre-trained models:

- Hugging Face Transformers - A library that gives access to many NLP models (BERT, RoBERTa, XLM-R, etc.) and LLM models (Phi4, Qwer, Llama). It can be used for classification, tokenization, embeddings, and more.
- BERTopic - A topic modeling tool based on embeddings (e.g., BERT). It is good for finding semantic topics in speeches and tracking how they change over time.
- spaCy-Transformers - Useful for basic text processing (tokenization, lemmatization)

3.3 Useful pretrained models for This Project

Below we listed some examples of pre-trained models for political speech analysis:

- Sentiment / emotions: Models like RoBERTa fine-tuned for sentiment analysis can be used to detect emotional tone in speeches (positive, negative, neutral).
- Propaganda detection: Fine-tuned transformer models based on ModernBERT-base (NCI Binary Propaganda Detector v2 and NCI Technique Classifier) can be used to classify speech fragments as propaganda and to identify specific propaganda techniques (e.g. name-calling, slogans, appeal to fear).
- Framing / metaphors: Token-classification models (BERT or XLM-R) can be fine-tuned to detect metaphors or narrative frames in text.
- Diachronic analysis: BERTopic with dynamic topic modeling can show how topics evolve over time in the speech corpus.
- Content understanding: Large language (decoder only) models, including Qwen, Llama, Phi, and GPT, which, together with Retrieval Augmented Generation, will be enabled to learn about Putin's speeches.

3.4 Datasets (Open Datasets)

For this project, we will use several open datasets suitable for political speech analysis:

- Putin Corpus (2012–2024) [6]:

  Corpus of Vladimir Putin's speeches from kremlin.ru. It is useful for analyzing topics, rhetoric, and propaganda.

- Chronorhetorics Corpus (until 2025) [5]

  A corpus designed for temporal rhetoric analysis — how politicians refer to past and future to legitimize power. Can be used as a comparison or for temporal analyses. Features many countries, including Russia.

4 Exploratory Data Analysis

4.1 Data Preprocessing - Putin Corpus dataset

The project is based mostly on a provided dataset in the form of a single `.json` file. This file contains a list of objects, where each object represents a single speech transcript. The key fields within each JSON object are:

- **`date`:** The date and time of the speech.
- **`transcript_unfiltered`:** The full, unfiltered transcript of the event.
- **`transcript_filtered`:** A cleaned version of the transcript, focusing on the core content of the speech.
- **`wordlist`:** A list of pre-processed (lemmatized) words from the transcript.
- **`title`:** The title of the event.
- **`kremlin_id`:** The document's identifier from the source system.
- **`place`:** The location where the speech took place.
- **`persons`, `teasers`, `tags`:** Additional metadata describing the document.

The dataset was loaded from a JSON file into a DataFrame. Missing values in the text columns were filled and converted to strings. The date column was converted to datetime, and rows with missing dates were removed. Only speeches containing "Putin" before a colon were retained, and the speaker prefix was removed. Year and month features were extracted. The main text was selected from the filtered transcript or, if missing, from the unfiltered transcript. Text was split into sentences and tokenized. Tokens were lowercased, filtered to keep only alphabetic words, stopwords were removed, and lemmatization was applied. Additional statistics were computed, including token counts, unique tokens, lexical diversity, and average sentence length. Rhetorical features such as the number of exclamation marks, question marks, and ellipses were also extracted.

4.2 EDA - Putin Corpus dataset

Basic metrics:

- **Total Number of Speeches:** 7629
- **Average Speech Length:** 921.04
- **Average Lexical Diversity:** 0.47
- **Average Word Count:** 921.04
- **Minimum Word Count:** 6
- **Maximum Word Count:** 26634

More analysis with plots:

The chart (Figure 1) shows the number of speeches delivered by Putin each year. The data were grouped by year and displayed as bars. Additionally, his presidential terms are highlighted: 2000-2008 in red and 2012-2025 in green, making it easier to visually relate speech activity to the periods he held office.

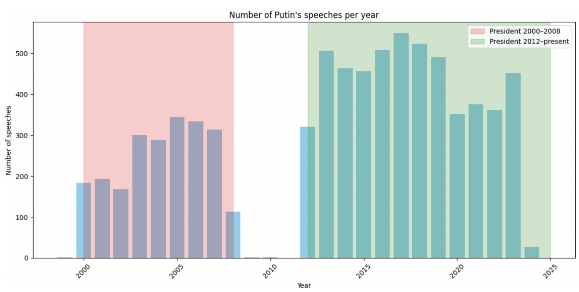


Figure 1: Distribution of speeches over the years

We extracted all bigrams (pairs of consecutive words) from the speeches after removing stopwords. To focus on meaningful content, we filtered out bigrams containing "weak" words such as would, like, or know. The top 20 most frequent bigrams were then identified and visualized as a word cloud (Figure 2), highlighting the phrases that appear most often in Putin's speeches.

Figure 2: Most common bigrams

We analyzed the tags associated with each speech to identify the most frequent tag per year (Figure 3).

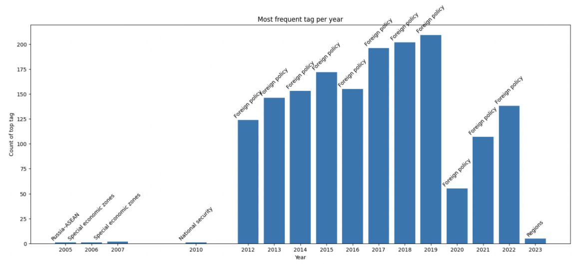


Figure 3: Most frequent tags per year

4.3 Data Preprocessing - Chronorhetorics-corpus

This Chronorhetorics-corpus dataset has a similar schema and JSON structure to the previous dataset, so we performed the same processing on it. The most important metadata are:

- **title:** For example, "Russian-Chinese talks have confirmed the strategic nature of the two countries' bilateral relations".
- **language:** Mostly English.
- **date:** In the format "2007-03-26".
- **location:** For example, "Moscow".
- **speaker:** For example, "KREMLIN" / "Putin".

- **`text`:** For example, “The Russian and Chinese leaders met six times in 2006...”.

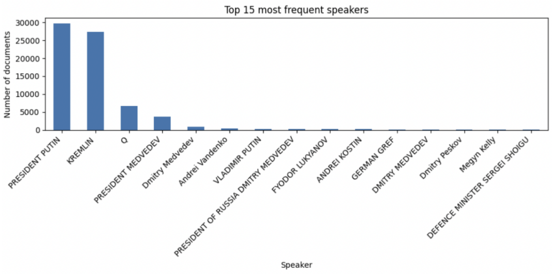


Figure 4: Different Speakers distrubution

On the (Figure 4) we can see that mostly there are two speakers with the biggest number of talks – Putin and Kremlin. We can also observe different names for Putin as in the pre processing we will ensure all of the Putin’s speeches are combined to the Putin person.

4.4 EDA - Chronorhetorics-corpus

The dataset of Vladimir Putin’s talks in this dataset contains:

- **Documents analyzed:** 33,129
- **Average speech length:** 486 words
- **Median speech length:** 140 words

In addition, the collection includes statements of a very diverse nature: from short, technical answers to long ideological narratives. In most cases we can observe the texts were a maximum of 250 words long, so they were rather short speeches (Figure 5).

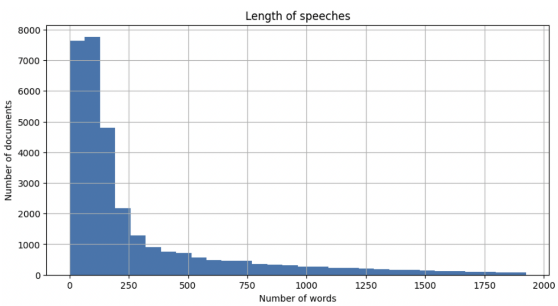


Figure 5: Speeches distribution over time

On the (Figure 6) we can observe the distribution of speeches between 1999 and 2023. We can see that between 2008 and 2010 Putin spoke a little more than in other years.

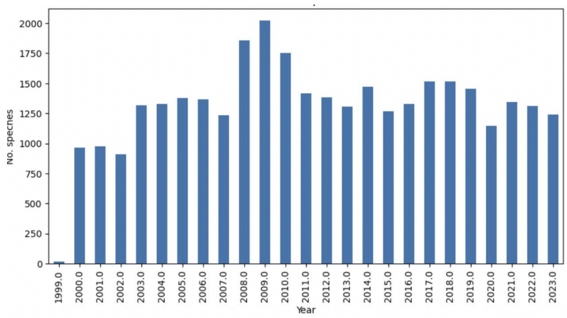


Figure 6: Speeches distribution over time

Based on the two frequency charts (single words (Figure 8) and bigrams (Figure 9)), we can describe the main topics of Vladimir Putin’s speeches in the following way.

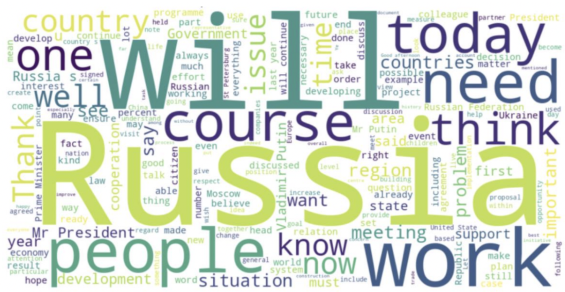


Figure 7: Speeches distribution over time

The most frequent single words clearly show that Putin often talked about Russia and the Russian state. Words like “Russia”, “Russian”, “country”, and “state” appear very often, which suggests a strong focus on national identity and sovereignty.

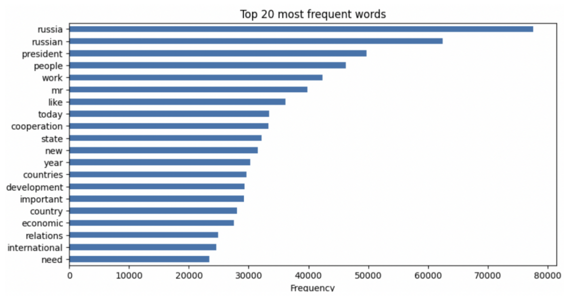


Figure 8: Speeches distribution over time

Another important group of words refers to people and society, such as “people”, “work”, “development”, and “important”. This indicates that many speeches addressed social issues, economic development, and the role of citizens. There is also a strong presence of political and institutional language. Words like “president”, “relations”, “international”, and “economic” may show that foreign policy and international cooperation were frequent topics. Time-related words such as "today" and "year" suggest that speeches were often connected to current events and annual summaries.

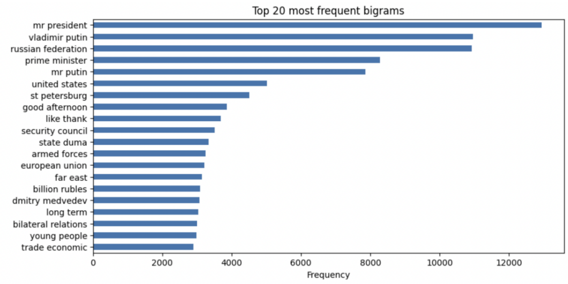


Figure 9: Speeches distribution over time

The bigrams give more concrete context. Bigram such as "Russian Federation", "state duma", and "security council" show that he often spoke within official political settings. International relations are also visible through phrases like "United States", "European Union", and "bilateral relations". In addition, some bigrams point to economic and regional topics, for example "billion rubles", "far east", and "long term". Overall, the charts show that Putin most often spoke about Russia, state power, international relations, economic development, and the role of people. His language is formal, institutional, and focused on governance, stability, and national interests.

5 Large Language Models for Putin's Talks Analysis

In this chapter, we describe the pipeline designed to analyze Vladimir Putin's speeches using Large Language Models (LLMs). The workflow consists of four main stages:

1. **Data Processing:** We utilized the `Chronorhetorics-corpus` (specifically the `ru-putin-speeches` folder). The dataset was filtered to retain only the speeches and responses given by Vladimir Putin.

2. **RAG Preparation:** We implemented a Retrieval-Augmented Generation (RAG) system using the FAISS library to store text embeddings and retrieve relevant context.

3. **Inference:** The selected LLM searches for knowledge in the text chunks and generates an answer based on the retrieved information.

4. **Evaluation:** We compared the performance of four models: `TinyLlama-1B`, `Qwen2.5-3B-Instruct`, `Phi-3.5-mini-instruct-4B`, and `GPT-4o`.

5.1 Retrieval-Augmented Generation (RAG) Setup

To allow the models to answer specific questions based on facts, we generated a vector database. The first step in understanding how models process this data is tokenization, where raw text is broken down into smaller units. Figure 10 illustrates this process using the example sentence: *"Good afternoon, colleagues. As previously agreed, we are going to discuss two issues today. The first one is related to ensuring security."* As shown in the visualization, the tokenizer splits the text into specific tokens, handling common words as single units and breaking down more complex words into sub-tokens (e.g., handling prefixes or suffixes with "#") to optimize processing.

Figure 10: Token splitting visualization

The process of building the knowledge base involved several key technical decisions. To convert the text into vector representations, we utilized the `all-MiniLM-L6-v2` embedding model. This specific model was chosen primarily for its high efficiency and inference speed, which allowed for rapid processing of the large dataset.

For the structure of the database, we applied a specific chunking strategy to improve retrieval accuracy. Instead of indexing whole documents, the text was split into smaller segments using a sliding window method. We set the chunk size to 800 characters with a step of 700 characters. This configuration creates a deliberate overlap between consecutive chunks, ensuring that important context is not lost at the boundaries of the segments. Finally, the generated embeddings were stored in a FAISS index using the `IndexFlatL2` metric. This metric measures the Euclidean distance between vectors, allowing the system to quickly find the most semantically similar content when queried.

5.2 LLM Integration and Inference Strategy

We evaluated three LLMs loaded with 4-bit NF4 quantization to optimize memory usage and gpt4o API. The FAISS index contained 82,266 vectors, the visualization of the entire vector database with Putin's statements was performed using t-sne and can be seen at Figure 11.

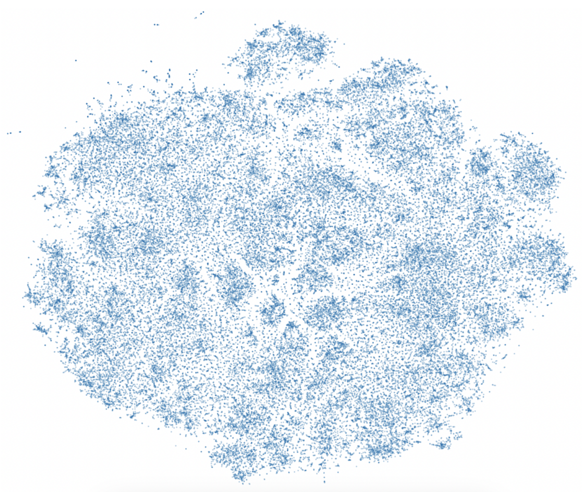


Figure 11: Vector database containing whole Putin's speeches database visualized using t-sne

In order for LLM to find the correct context in this database, we used an algorithm in which we match the k semantically closest chunks for each question. The inference pipeline retrieved the top k semantic chunks from the FAISS index, which were then injected into a system prompt explicitly restricting the model to answer solely based only on the provided context. To analyze performance variations, we tested retrieval depths of $k = 3$ and $k = 10$, combined with generation temperatures of 0.1 (deterministic) and 0.7 (creative).

6 LLMs Evaluation

The evaluation process included three compact, open-weights models: `TinyLlama-v1-1B`, `Qwen2.5-3B-Instruct`, `Phi-3.5-mini-instruct-4B` and one much larger model `GPT-4o`.

To benchmark their effectiveness and hallucinations we adopted an "LLM-as-a-Judge" evaluation protocol. Instead of relying solely on manual human annotation, we utilized `Gemini 3 Pro` as an automated evaluator. This large-scale model analyzed the responses generated by the candidate models to determine their quality and consistency. The evaluation focused on two primary metrics:

- **Effectiveness:** The judge assessed whether the model provided a correct, coherent, and relevant answer.
- **Answer Refusal:** We measured how effectively the models recognized insufficient data and refused to answer (e.g., stating "I don't know") rather than hallucinating incorrect facts.

6.1 Answer Refusal Comparison

Figure 12 illustrates the overall refusal rate. `GPT-4o` demonstrated the highest adherence to safety constraints with a refusal rate of 0.43, followed closely by `Phi-3.5-mini` (0.39). In contrast, `TinyLlama-1B` failed to refuse any queries (0.00), suggesting a high tendency to generate hallucinations regardless of the available context.

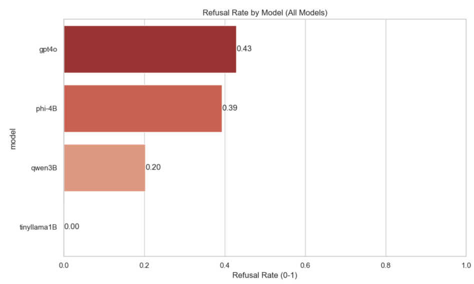


Figure 12: Overall refusal rate by model. Larger models show higher capabilities in recognizing insufficient context.

We also analyzed which types of questions were most difficult for the system to answer (Figure 13). Questions related to "Change over time" resulted in the highest refusal rate (0.51). This highlights the limitation of the RAG approach based on static chunks, which struggles to aggregate temporal trends spanning multiple years.

Finally, Figure 14 provides a granular breakdown. `GPT-4o` (dark red) consistently exhibits the highest caution, particularly in "Date range" and "Metaphors" categories, where it refused over 80% and 60% of queries respectively. Smaller models, especially `TinyLlama` and `Qwen`, often

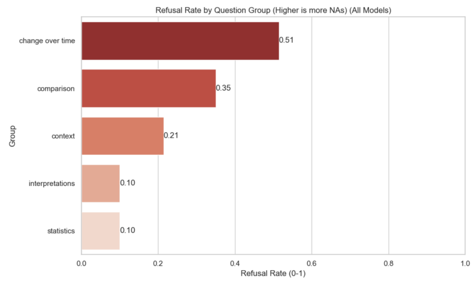


Figure 13: Refusal rate aggregated by question category.

attempted to answer these complex queries, likely resulting in lower factual accuracy.

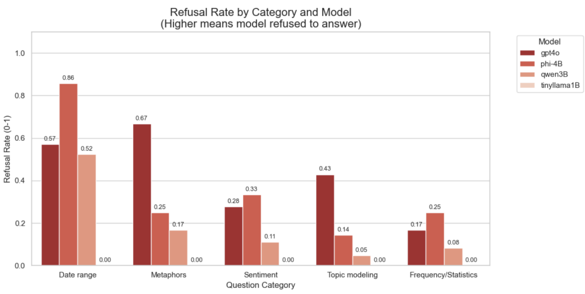


Figure 14: Detailed refusal rates by specific question types and models.

6.2 Answer Effectiveness Comparison

After analyzing the refusal rates, we evaluated the quality of the provided answers using a larger model to verify consistency with ground truth data. We employed a 1–5 scale defined as follows: 1 – complete hallucination; 2 – partial hallucination; 3 – incorrect context; 4 – correct but slightly misaligned with user intent; 5 – comprehensive and perfect response. Finally, all scores were normalized to the range $[0, 1]$ by dividing the raw score by 5.

Figure 15 presents the aggregated performance. Unsurprisingly, the largest model, `GPT-4o`, achieved the highest effectiveness score of 0.82. However, the most notable result is the performance of `Phi-3.5-mini-instruct-4B`. Despite having significantly fewer parameters than GPT-4o, it achieved a score of 0.71, clearly outperforming `Qwen2.5-3B` (0.58) and doubling the performance of `TinyLlama-1B` (0.33).

The detailed breakdown in Figure 16 and the heatmap in Figure 17 highlight specific strengths and weaknesses. `Phi-4B` proved exceptionally capable in analytical categories such as "Topic modeling" (0.85) and "Sentiment" (0.82), where

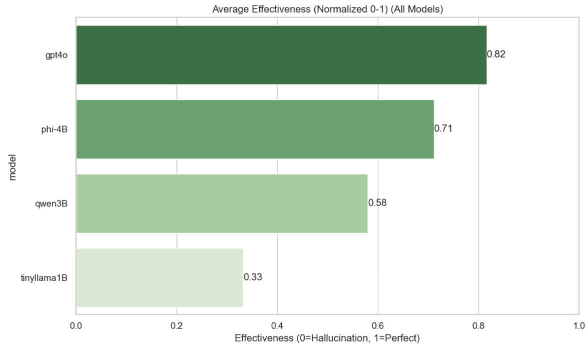


Figure 15: Overall effectiveness rate by model.

it nearly matched the proprietary GPT-4o model. In contrast, abstract tasks like "Metaphors" were too complex for the smallest model, `TinyLlama`, which failed to produce any valid answers in this category (0.00).

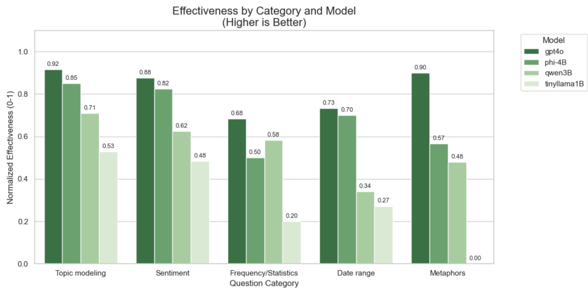


Figure 16: Detailed effectiveness rates by specific question types.

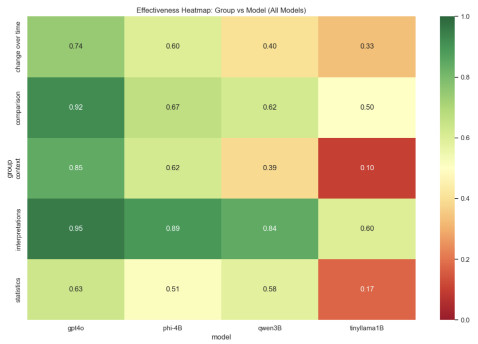


Figure 17: Heatmap of average effectiveness per category. Darker green indicates higher accuracy.

Finally, Figure 17 aggregates all of the models by the task group. The "Interpretations" group was the most consistently handled category across all models, with even smaller architectures scoring reasonably well. Conversely, questions requiring strict contextual precision ("Context") or handling specific data points ("Statistics") caused significant degradation in the performance of `TinyLlama` and `Qwen`.

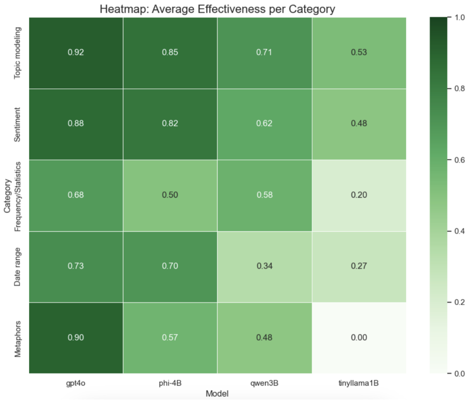


Figure 18: Heatmap comparing question group and model performance.

Ultimately, it turned out that the models performed worst overall with the "Data range" and "statistics/frequency" question types. In these types of tasks, questions often concerned, for example, changes in behavior over the entire period from 2001 to 2024 or counting occurrences related to the military. These tasks are particularly difficult for LLMs that rely on only the 10 closest chunks, which may be insufficient to fully answer the questions. LLM models performed poorly on average in the metaphor category.

On the other hand, LLM models performed excellently on tasks such as topic modeling and sentiment analysis. Phi and GPT models reign supreme here, achieving results above 80% effectiveness. The best group of questions for LLM models are those related to interpretation, where all three models achieved very high scores. They also perform well on the comparison task group.

6.3 Qualitative Analysis of RAG Parameters

Beyond quantitative metrics, we observed several distinct behavioral patterns related to model size, context window (k), and generation temperature. The most significant qualitative findings are summarized below:

- **Small Model Limitations ("The Void Effect"):** The `TinyLlama-1.1B` model exhibited severe "Generative Confabulation." When the retrieved chunks did not contain the answer, the model refused to admit ignorance. Instead, it showed the "Void Effect," filling data gaps with hallucinations based on its pre-training. Additionally, it suffered from "Loss of Persona," frequently forgetting it was analyzing Putin's specific rhetoric and shifting to a generic Western geopolitical perspective.

- **Context Contamination ("The Soup Effect"):** Increasing the number of retrieved chunks ($k = 10$) for mid-sized models like `Qwen` often proved detrimental, leading to "Context Contamination.". For instance, the model syntactically merged a description of place in St. Petersburg with a political text about Poland, creating a false narrative that placed Russian landmarks within Polish borders.

- **Temperature and Temporal Bias ("Time Freeze"):** Lowering the generation temperature to 0.1 successfully eliminated poetic hallucinations and style fabrications found at higher temperatures. However, this stricter setting revealed a critical "Time Freeze" flaw: the models treated historical data (e.g., from 2000) as absolute, current truth, rendering them unable to weigh it against newer information present in the context.

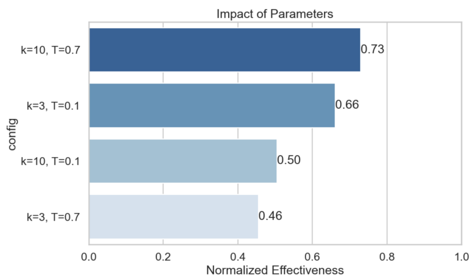


Figure 19: Comparison on different parameters

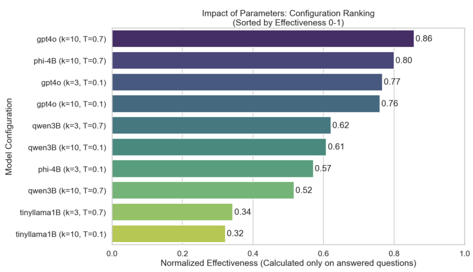


Figure 20: Overall effectiveness score

Model	Question (Topic)	Generated Response (Summary)	Type of Error	More specific large LLM judgement
TinyLlama	China (Friendship)	"In a 2016 interview with the BBC Putin said: Our friendship is based on shared values: democracy, freedom, and human rights."	**Total Fabrication** (Political inversion).	Putin has not claimed shared "democratic values" with China during BBC interview in 2016. He has never mentioned that.
TinyLlama	Poland (Metaphors)	"Poland is a shield against the cold winds of history," "A bridge over troubled waters."	**Poetic Hallucination** (AI invention).	Pure poetry invented by the AI. Putin uses language of hostility or historical grievance toward Poland, never poetic appreciation.
Qwen 3B	Poland (Metaphors)	"Statue of Pushkin" and "Palace Bridge".	**Context Contamination** (Merging unrelated chunks).	Context Contamination. The model attributed landmarks from St. Petersburg (Russia) to Poland simply because they appeared in the same batch of text chunks.
TinyLlama	USA (Metaphors)	"The US is like a giant pearl slowly turning into a rock," "A sleeping bear."	**Poetic Hallucination** (AI invention).	Non-existent metaphors. "The Bear" is a symbol universally associated with Russia, not the USA.
TinyLlama	Ukraine (Metaphors)	"Ukraine is a pearl in the ocean," "A rose in the garden."	**Poetic Hallucination** (AI invention).	Naive, positive imagery that contradicts the Kremlin's aggressive war rhetoric regarding Ukraine.
Qwen 3B	Poland (Friend/Enemy)	"In Putin speeches Poland is presented as a partner or ally."	**Total Fabrication** (Political inversion)	**Critical Geopolitical Error.** In Putin's speeches, Poland is consistently portrayed as a historical adversary or a vassal of the West, never an ally.
TinyLlama	Poland (Friend/Enemy)	"Poland appears as a friendly and valued partner of the United States."	**Wrong perspective**	**Persona Loss.** The model answered from the perspective of the US State Department, not based on Putin's speeches.
TinyLlama	Ukraine (Justification)	"The Soviet invasion of Afghanistan (1979)."	**Time Freeze** (Using 1979 data only).	**Logical Failure.** The model used an event from 40 years prior (Afghanistan) to justify current actions in Ukraine, which holds no logical connection.
Qwen 3B	Sovereignty (Timeline)	"It became a main topic two days ago."	**Time Freeze** (think of data as current state)	**Relative Time Error.** The model treated the phrase "I said two days ago" found in the text as a historical absolute date.
Qwen 3B	Germany (2003 vs 2014)	"In 2014, relations deepened and strengthened."	**Total Fabrication** (Political inversion	**Factual Error.** 2014 marked the collapse of relations due to Crimea. The model applied a positive sentiment from an earlier chunk to a later date.
Qwen 3B	USA	"Putin describes USA as a possible partner."	**Time Freeze** (Using 2001 data only).	**Logical Error:** LLM found most suited chunk from 2001 and didn't look at other dates

Table 1: The most absurd and factually dangerous responses generated during the test.

7 BERTopic

We decided to share with all our colleagues dashboard that allows for interactive search of topics in Putin speeches. The dashboard is available on website michalpuscian.com

In these section we will present our methodology regarding BERTopic and answering some of the posed questions with it.

7.1 Description

BERTopic is a topic modeling technique that leverages huggingface transformers and c-TF-IDF to create dense clusters allowing for easily interpretable topics whilst keeping important words in the topic descriptions.

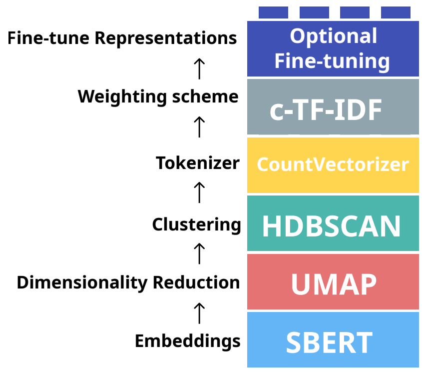


Figure 21: Diagram explaining the stages of detecting topics with BERTopic

BERTopic consists of 5 standard stages (Figure 21) and one optional. Although these steps are the default, there is some modularity to BERTopic. All stages are somewhat independent from one another. For example, the tokenization step is not directly influenced by the embedding model that was used to convert the documents which allow us to be creative in how we perform the tokenization step. As a result, BERTopic is quite modular and can maintain its quality of topic generation throughout a variety of sub-models. In other words, BERTopic essentially allows you to build your own topic model which we can see in Figure 22

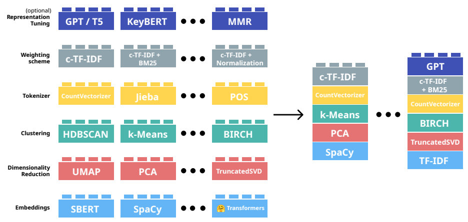


Figure 22: Modularity of BERTopic. We can use custom algorithm for each step thus creating our own topic modeling algorithm.

7.2 Our BERTopic

7.2.1 Data preprocessing

We divide each Putin speech into 5 sentences. The goal is to obtain more granular topics than would result from treating entire speeches as single documents. For this we use spacy en-core-web-sm. For each chunks we save the timestamp for analysis of topics evolution over time.

7.3 Customized stages

For embeddings we use the default Sentence-Transformer with ("all-MiniLM-L6-v2") provided by BERTopic. We perform dimensionality reduction with UMAP with cosine metric, 15 neighbors, 5 components and 0 minimum distance. The clustering is performed by HDBSCAN with minimum cluster size of 10 and euclidean metric. For tokenizer we use CountVectorizer with ngram range of 1 to 3 which allows for more accurate ngrams topics representations, we delete the english stop words and select minimum frequency of word to qualify for topic representation to 2.

For topic representation we employ a two-stage pipeline combining KeyBERTInspired() and MaximalMarginalRelevance(diversity=0.3), where KeyBERT extracts candidate keywords based on semantic similarity to document embeddings, and MMR subsequently diversifies the final representation to reduce redundancy among topic terms. We modify the default c-TF-IDF weighting scheme by using ClassTfidfTransformer with BM25 weighting enabled and frequent word reduction, which helps to down-weight common terms that appear across multiple topics and produces more discriminative topic representations. To guide the topic discovery process, we define a seed topic list containing six predefined topic seeds covering key themes expected in the corpus such as sovereignty, military conflict, Poland, Germany, and Ukraine-related narratives. These seeds provide soft guidance to HDBSCAN clustering, encouraging the model to discover topics aligned with our analytical framework while still allowing for emergence of unexpected themes. Finally, we fit the model on the preprocessed speech chunks and obtain both hard topic assignments and probability distributions over topics for each document.

7.4 Dynamic topic modeling

We use dynamic topic modeling (DTM) to obtain evolutions of topics over time.

DTM is a collection of techniques aimed at analyzing the evolution of topics over time. These methods allow you to understand how a topic is represented across different times. For example, in 1995 people may talk differently about environmental awareness than those in 2015. Although the topic itself remains the same, environmental awareness, the exact representation of that topic might differ.

BERTopic performs DTM by calculating c-TF-IDF representation for each topic and timestep. Next, there are two main ways to further fine-tune these specific topic representations, namely globally and evolutionary.

A topic representation at timestep t can be fine-tuned globally by averaging its c-TF-IDF representation with that of the global representation. This allows each topic representation to move slightly towards the global representation whilst still keeping some of its specific words.

A topic representation at timestep t can be fine-tuned evolutionary by averaging its c-TF-IDF representation with that of the c-TF-IDF representation at timestep t-1. This is done for each topic representation allowing for the representations to evolve over time.

We perform both fine-tunings of topics representations.

7.4.1 Possible modifications

We could create overlapping chunks instead of non-overlapping ones. This could result in better topic representation because of information flow continuity.

7.5 BERTopic results

We obtain 500 topics together with their names, representations (list of words) and speeches that are representative to each topic. We also perform DTM to get evolutions of topics over time.

7.5.1 Answering questions

With these results we are ready to answer these questions:

1. In references to Poland, are historical or contemporary contexts more frequent?
2. Provide quotes where he describes Poland in historical terms
3. When do themes of “sovereignty” begin to dominate foreign policy discourse?
4. What three different arguments does he invoke when speaking about sanctions?
5. What historical events does he use to legitimize actions toward Ukraine?
6. What are Putin’s most common arguments for strengthening the army?

For each question we extracted topics similar to the question in terms of cosine similarity. We inspected representative topics to make sure they are correctly extracted.

Question: In references to Poland, are historical or contemporary contexts more frequent? and Provide quotes where he describes Poland in historical terms.

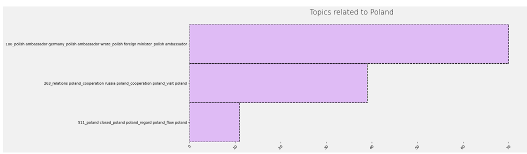


Figure 23: Topic Modeling concerning Poland

Three topics were obtained with their counts visible on Figure 23. The most mentions about Poland were related to historical World War II-era events, specifically diplomatic exchanges between Polish and German officials in 1938–1939, including discussions of territorial arrangements regarding Czechoslovakia and antisemitic policies (Topic 186, n=70). Example of representative document is: *”On September 20, 1938, Polish Ambassador to Germany Józef Lipski reported to Minister of Foreign Affairs of Poland Józef Beck on the following assurances made by Hitler...”*. Contemporary contexts appeared in two smaller topics: diplomatic relations and cooperation between Russia and Poland (Topic 263, n=39) and energy policy disputes, particularly regarding gas transit and sanctions (Topic 511, n=11). Overall, historical references (n=70) outweigh contemporary ones (n=50 combined), suggesting Putin more frequently invokes Poland’s WWII-era history often to draw controversial parallels, than he discusses current bilateral relations.

Question: When do themes of ”sovereignty” begin to dominate foreign policy discourse? For this question, we employ Dynamic Topic Modeling (DTM) to trace the temporal evolution of sovereignty-related discourse. Three relevant topics were identified: Topic 40 (n=255) concerning Iraq and international law, Topic 203 (n=63) relating to traditional political sovereignty, and Topic 292 (n=32) addressing technological sovereignty.

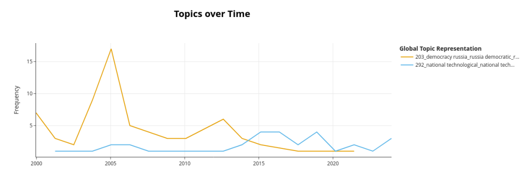


Figure 24: Frequency of two topics related to sovereignty over the years

As visible in Figure 24, the narrative undergoes a gradual but distinct transformation. In earlier periods, sovereignty discourse centered on *classical Westphalian principles* defending Iraq’s sovereignty against Western intervention and emphasizing that military operations *”ran counter to global public opinion, the principles and norms of international law and the UN Charter”*. Concurrently, Topic 203 reflects domestic sovereignty framing, with Putin declaring that *”the Russian Constitution is one of the most democratic in the world”* and celebrating *”the Day of Russian Sovereignty”*. However, in later periods, particularly post-2014, the discourse shifts toward *technological sovereignty* (Topic 292), with emphasis on the *”National Technology Initiative”* and *”national technological sovereignty projects”*. Putin argues that *”anyone who wants to take the lead in the world today has to put their focus on [...] developing technology and education”*. This evolution reflects a broader reconceptualization: from sovereignty as protection against external political interference to sovereignty as self-sufficiency in strategic technologies, a shift likely accelerated by Western sanctions and growing technological competition.

Question: What three different arguments

does he invoke when speaking about sanctions? Four topics were obtained with their counts visible on Figure 25.

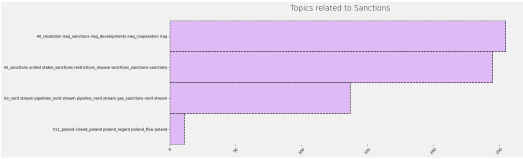


Figure 25: Topic Modeling concerning sanctions

Three distinct argumentative strategies emerge from the analysis. First, Putin frames sanctions as *ineffective and futile*, stating: *"I have no doubt that this, as lawyers say, is an exercise in futility, and nothing will come of it"* (Topic 42, n=245). Second, he employs a *boomerang argument*, claiming sanctions harm their initiators more than Russia: *"This policy primarily hurts those who initiated it"* and *"imposing sanctions always causes some damage, including for those who impose them"*. Third, he invokes *collective damage to global stability*, arguing that *"the situation with the so-called sanctions is damaging for the global economy [...] and it is damaging for the Russian-European relations"*. Additionally, in the context of Iraq (Topic 40, n=255), Putin frames Western sanctions and military interventions as violations of *"the principles and norms of international law and the UN Charter"*, positioning Russia as a defender of multilateral order. The Nord Stream-related topic (Topic 93, n=137) reflects how sanctions on energy infrastructure became a focal point of this rhetoric.

Question: What historical events does he use to legitimize actions toward Ukraine? Eleven topics were obtained with their counts visible on Figure 26.

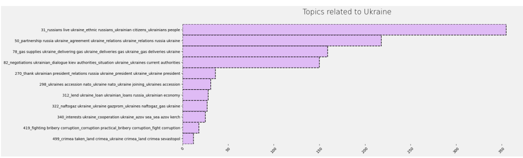


Figure 26: Topic Modeling concerning Ukraine

Putin employs several historical legitimization strategies. First, he invokes *Ukraine's own secession from the USSR* as a precedent for Crimean self-determination: *"When Ukraine seceded from the USSR it did exactly the same thing"* (Topic 499, n=12). Second, he references *European precedents of ethnic minority protection*, comparing Russia's actions to Hungary and Romania issuing passports to co-ethnics abroad: *"Hungary and Romania went as far as give away passports to ethnic Hungarians and Romanians [...] are ethnic Russians living in Ukraine worse than Romanians, Poles or Hungarians?"* (Topic 31, n=355). Third, he cites *Ukraine's original Declaration of Independence*, which stated Ukraine would be a neutral state, to delegitimize NATO aspirations: *"When Ukraine gained independence, the Declaration of Independence explicitly stated that Ukraine was a neutral state"* (Topic 298, n=31). Fourth, he constructs a narrative of *spiritual-historical belonging*, framing Crimea as *"the spiritual source of the development of a multifaceted but solid Russian nation"*. Finally, he references the *UN Charter's right to self-determination* to frame the Crimean referendum as internationally legitimate, while simultaneously invoking Russia's post-Soviet generosity: *"Not only did Russia recognise these countries, but helped its CIS partners"* (Topic 312, n=28).

Question: What are Putin's most common arguments for strengthening the army? Five topics were obtained with their counts visible on Figure 27.

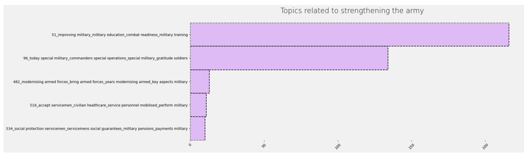


Figure 27: Topic Modeling concerning Army

Putin employs several distinct argumentative strategies. First, he emphasizes *strategic nuclear deterrence and technological modernization*, arguing that *"the combat capability of the nuclear triad [...] directly depend on the stability, effectiveness and reliability of these systems"* (Topic 51, n=216). Second, he invokes *historical military lessons*, particularly from WWII, warning against repeating past mistakes: *"the students in military academies were learning about trench warfare, and when the tank offensives began [...] the situation was markedly not in our favour"*. Third, he frames modernization as a response to *global instability*: *"the world situation has not become more stable in the past few years. It has even deteriorated in some regions"* (Topic 482, n=13). Fourth, he constructs a *patriotic-heroic narrative*, comparing contemporary soldiers to *"the heroes of the War of 1812, the First World War or the Great Patriotic War"* (Topic 96, n=134). Fifth, he argues for *social investment in military personnel* as essential to institutional strength: *"it is only by listening carefully to the military's pressing problems and solving them that we can bolster the social status of defenders of the Motherland"* (Topics 516 and 534, n=21 combined). This includes concrete measures such as tripling service pay and improving military pensions to *"enhance the prestige of military service in society"*.

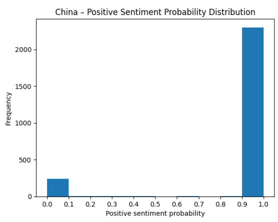
8 Sentiment Analysis Toward Countries and Topics

8.1 Method: sentence filtering and RoBERTa aggregation

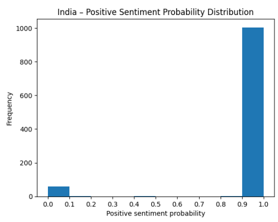
We estimated the sentiment of Putin's rhetoric toward specific actors by restricting the analysis to sentences explicitly mentioning the country/institution in question (via curated keyword lists). Each retained sentence was scored with a RoBERTa sentiment classifier. For each actor and year, we then computed the average sentiment (mean score over all actor-mentioning sentences in that year), yielding a yearly sentiment trajectory.

8.2 Country-level findings and event alignment

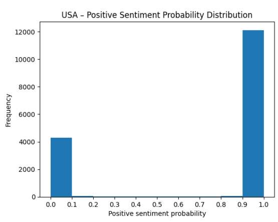
Figure 28 shows that sentiment is more positive toward China and India than toward the USA and the European Union.



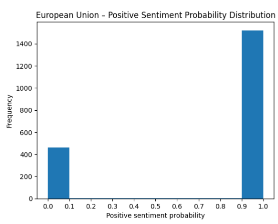
(a) China



(b) India



(c) USA



(d) European Union

Figure 28: Yearly average RoBERTa sentiment computed over sentences mentioning each actor (filtered by actor keywords).

Poland is discussed with a rather negative tone (Figure 29). Ukraine and USA follow a similar evolution over time, consistent with Kremlin discourse linking Ukraine-related developments to the USA's long-term support for Ukraine. The main variations follow major events:

- 2004: Orange Revolution
- 2010: Yanukovych elected
- 2014: Crimea / Euromaidan
- 2022: Full-scale invasion

Notably, an increase in sentiment approximately two years prior to the 2022 invasion is visible in the Ukraine/USA trajectories (Figure 30), which may indicate rhetorical appeasement preceding escalation.

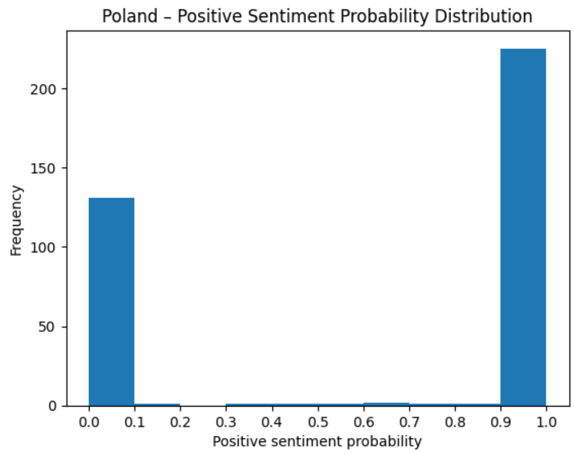


Figure 29: Yearly sentiment toward Poland, showing an overall negative tone compared to other actors.

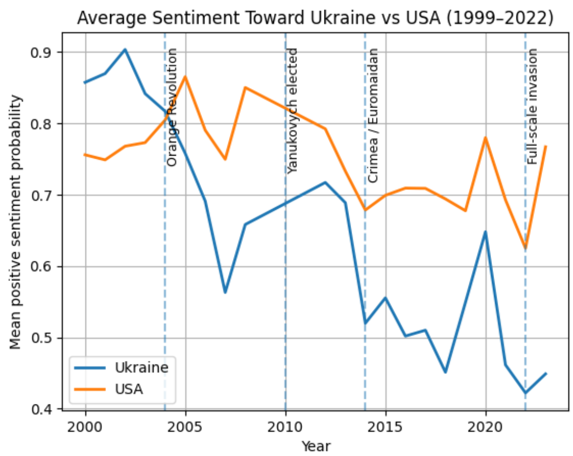


Figure 30: Sentiment comparison over twenty years. Ukraine and USA show similar temporal evolution; shifts align with major events (2004, 2010, 2014, 2022).

8.3 Topic-level sentiment

Beyond countries, we computed topic-level sentiment by grouping sentences by topic and averaging RoBERTa scores within each topic bucket. Figure 31 shows that agriculture- and Ukraine-related topics are low in sentiment, whereas cooperation-related topics are high.

Figure 31: Topics ordered by average sentiment. Agriculture and Ukraine-related topics are low; cooperation-related topics are high.

9 Word frequency analysis

To address statistical questions, word frequency analysis was applied alongside custom lexicons to identify mentions of specific concepts. Occurrences of individual words, such as economic terms, were counted, as well as co-occurrences of words from different lexicons within the same sentence, for example "friendship" and "China." These analyses were then compared with the results produced by the least effective LLM model on statistics tasks, TinyLlama1B. All analyses were conducted on the Putin Corpus dataset. For the question

Which three economy-related terms appear most often?, the analysis indicates that development, cooperation, and interest are the most frequent economic terms in the dataset, as illustrated in Figure 32. In contrast, TinyLlama1B produced the answer: *"According to the given text, Putin uses the following three economy-related words most often: 1. Strengthening 2. Property 3. Tax"* which was confirmed to be incorrect.

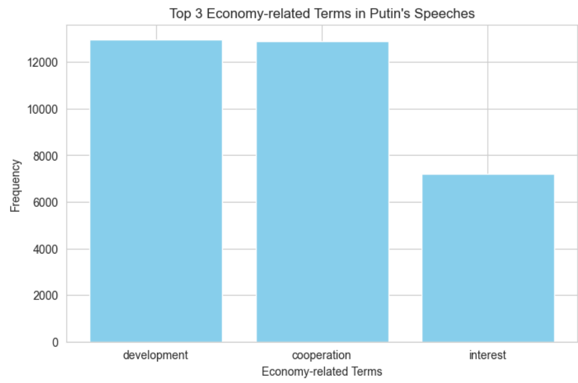


Figure 32: Three most frequent economy-related words.

To answer the question **How many times does the word "modernization" appear in a military context?**, a plot (Figure 33) was created to show its frequency over time, revealing a peak in 2012. TinyLlama1B, instead of providing only a numeric answer, also offered examples: *The word "modernization" appears multiple times in Vladimir Putin's military context speeches. Here are some examples: In his address to the Russian Armed Forces on May 20, 2019, Putin mentioned that the Russian military should modernize its equipment and infrastructure to meet the challenges posed by modern warfare [...]"*.

Figure 33: Modernization mentions in military context over time.

Similar plots (Figures 34 and 35) were produced for the questions **How often does he speak about "friendship" in relation to China?** and **In which years do the most references to World War II occur?** This approach allows the frequency and temporal patterns of mentions to be easily observed. For the first question, TinyLlama1B correctly identified the frequency and additionally provided context, describing how China was referred to in terms of friendship: *"Putin frequently mentions his friendship with China in public speeches and interviews. He has spoken about it several times during his presidency, including during his state visit to China in 2013, where he emphasized the importance of maintaining strong ties between the two countries. Putin has also praised China's economic growth and its role as a major player in global politics, highlighting the benefits of having a close relationship with China [...]"*.

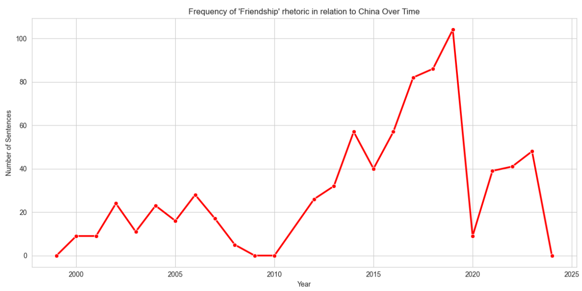


Figure 34: Friendship mentions related to China over time.

Regarding the final question, TinyLlama1B did not provide a specific count but noted the context in which World War II was referenced: *"Putin refers to World War II most often in the context of the tragedy of the September 11 attacks, specifically the rescue efforts by Russians who were near the World Trade Center [...]"*.

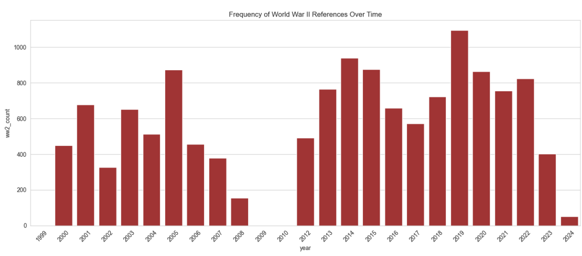


Figure 35: World War II mentions over time.

10 Propaganda detection

The entire corpus of Putin's speeches was analyzed to detect propaganda. After identifying propaganda content, only sentences classified as propaganda were retained for further analysis. These sentences were then used to examine how dominant topics and propaganda techniques changed over time.

Propaganda detection was performed using the NCI Binary Propaganda Detector v2, based on the ModernBERT-base model. The results (Figure 36) show that in all but one year, propaganda sentences outnumbered non-propaganda ones. This difference becomes especially pronounced between 2012 and 2024, corresponding to Putin's second presidential period. Overall, approximately 54% of the speeches in the dataset were classified as propaganda.

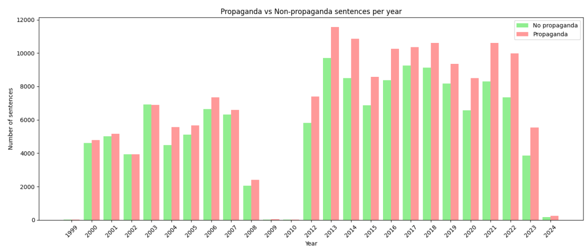


Figure 36: Number of propaganda and non-propaganda sentences over time.

10.1 Propaganda topic modeling

The next step focused on identifying dominant topics for each year (Figure 37). Topic modeling was performed using only propaganda sentences with a prediction score above 0.9. For each year, the dominant topic was determined by selecting the most frequently assigned topic label.

Sentence embeddings were generated using the all-MiniLM-L6-v2 model, and topic modeling was conducted with BERTopic. The results reveal clear temporal patterns. Ukraine-related topics emerge in 2014, corresponding to the Donbas conflict, and reappear in 2022 during the full-scale invasion. Election-related topics occur only in 2000, when Vladimir Putin first assumed the presidency. Pandemic-related topics dominate in 2020 and 2021, aligning with the COVID-19 crisis, while Syria-related topics first appear in 2015 following Russia's military intervention. Overall, the detected topics closely mirror major real-world political events.

10.2 Propaganda techniques

Finally, propaganda techniques were analyzed over time using the same set of propaganda sentences (Figure 38). For this purpose, the NCI

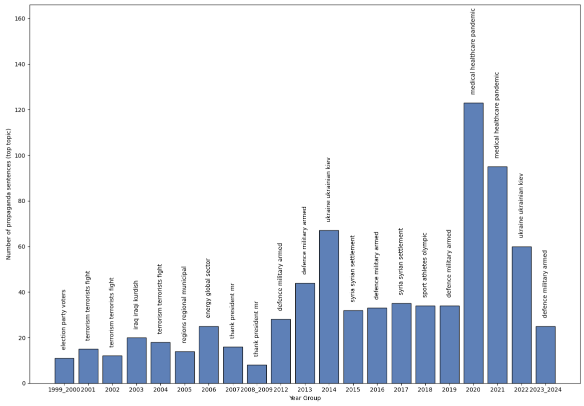


Figure 37: Dominant propaganda topic over time.

Technique Classifier, based on the ModernBERT-base model, was applied.

Across all years, the most dominant technique is name-calling and labeling. Other frequently used techniques include slogans, reductio ad Hitlerum, doubt, and appeals to fear or prejudice. Importantly, the overall distribution of techniques remains largely stable over time, indicating that while the topics of propaganda change, the rhetorical strategies used to convey them stay remarkably consistent.

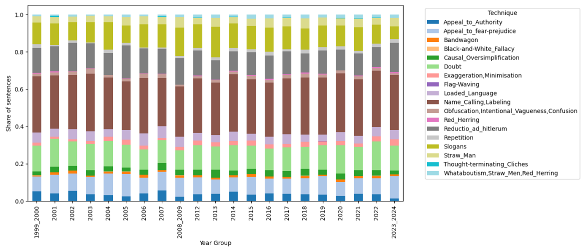


Figure 38: Dominant propaganda technique used over time.

10.3 Evolution of the ”Russian World” Narrative

To further refine the analysis of ideological shifts, the frequency of the concept of the ”Russian World” (*Russkiy Mir*) was examined. The methodology involved a hybrid approach: first, a regex-based search identified mentions of the phrase and its variations. Subsequently, a Named Entity Recognition (NER) pipeline, utilizing the `bert-large-finetuned-conll03-eng` model, was applied to extract the linguistic context and entities surrounding these mentions.

The results (Figure 39) indicate that the term was virtually absent from official rhetoric until 2006. A significant turning point occurred in 2007, marked by a sharp peak (11 mentions) coinciding with the formal establishment of the *Russkiy Mir Foundation*. Following a period of relative dormancy, the narrative resurfaced with consistent intensity between 2012 and 2014, and again reached a sustained plateau during the 2022–2023 invasion of Ukraine. This pattern suggests that the ”Russian World” is not a constant rhetorical element but a strategic tool activated during periods of institutional expansion or geopolitical conflict.

Figure 39: Frequency of ”Russian World” (Russkiy Mir) mentions in Putin’s speeches

11 Analysis of statements about Poland

In the previous section on sentiment analysis, we discovered that sentiment towards Poland is more negative than that towards the US, EU, China, and other mentioned countries. In this section, we examine more closely what kind of statements accompany Poland.

To categorize the rhetoric in the quetion **In what context does “Poland” most often appear? (enemy, partner, ally, neighbor)**, we employed a `zero-shot classification` approach using the `facebook/bart-large-mnli` model.

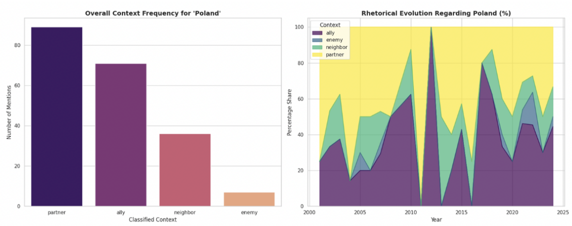


Figure 40: Overall context frequency (left) and rhetorical evolution regarding Poland (right).

This method allows for classifying text into predefined labels, leveraging the model’s internal knowledge to determine the most likely context for each sentence containing the word ”Poland” or ”Polish”. The results (Figure 40) highlight a significant shift in Russian official discourse:

- **Historical Dominance of Pragmatism:** Aggregated data shows that **partner** and **ally** are the most frequent labels. This is primarily driven by the rhetoric of the early 2000s, where Poland was often framed within the context of European cooperation.

- **The ”Enemy” Pivot:** While the **enemy** label is the least frequent overall, the temporal analysis reveals it only begins to occupy a visible share of the rhetoric after 2020, intensifying during the 2022–2024 period.

11.1 Structural Construction of the ”Enemy” Image

To answer the question **How does Putin construct the image of the “enemy”?**, we performed semantic clustering of hostile statements. We used *Sentence-BERT* for high-dimensional embeddings, *UMAP* for dimensionality reduction, and *DBSCAN* to identify distinct thematic landscapes. The analysis revealed five primary rhetorical clusters (Figure 41, 42):

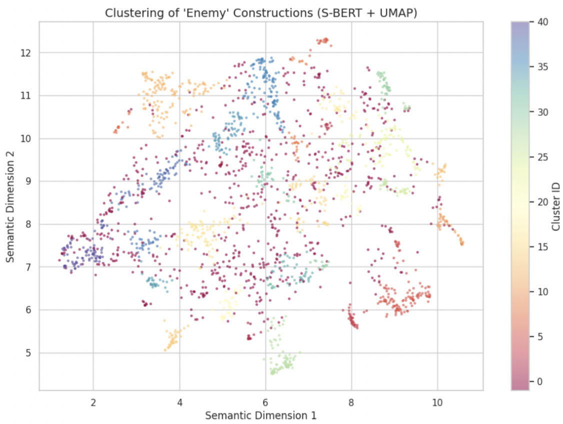


Figure 41: Clustering of ’Enemy’ constructions

- **Cluster 12 (Ukraine):** Focuses on Ukrainian sovereignty and the conflict as a primary site of confrontation.

- **Cluster 36 (NATO & Security):** Frames the enemy through the lens of Western military alliances and global security architecture.

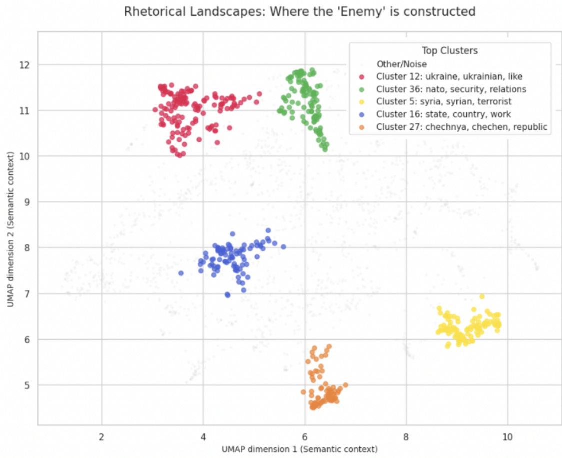


Figure 42: Top 5 clusters of the Putin’s Enemy Rhetoric

- **Cluster 5 (Syria & Terrorism):** Utilizes the rhetoric of international terrorism and external intervention.

- **Cluster 16 (Statehood/Work):** Relates to the internal stability of the state in opposition to external pressures.

- **Cluster 27 (Chechnya):** A historically rooted cluster regarding separatism and internal enemies.

11.1.1 Poland’s Position in the Enemy Narrative

By overlaying mentions of Poland onto these clusters, we identified how the ”Polish threat” is specifically constructed. The most significant results show that Poland is primarily embedded in Cluster 12 (Ukraine & Sovereignty), Cluster 36 (NATO & Security) and Cluster 35 (Nuclear, missiles & weapons)

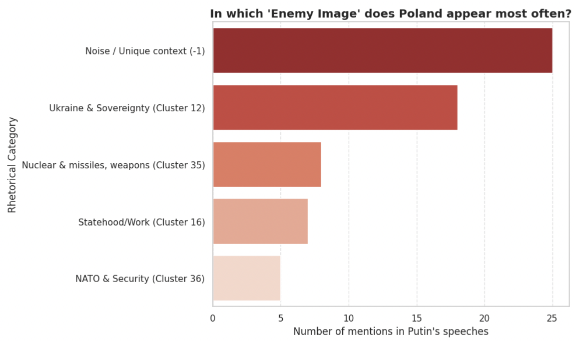


Figure 43: Top 5 clusters in which Poland appeared most often

Poland is rarely presented as an independent antagonist. Instead, its "enemy image" is functional and derivative: it is framed either as a primary instigator within the Ukrainian context or as a proxy for NATO expansion.

12 Conclusions

Our analysis of Vladimir Putin's speeches (1999–2025) reveals a clear shift in political narrative and provides a critical evaluation of modern NLP methodologies for political discourse.

12.1 Key Methodological Findings

- **LLM Limitations:** Large Language Models (LLMs) proved unreliable for **statistics** (frequent hallucinations), **metaphors** (poetic inventions), and **"change over time" analysis**. The RAG approach struggled to aggregate data across large date ranges, often suffering from "Time Freeze".
- **Success of BERT-based NLP:** Specialized models (`RoBERTa`, `BERTopic`) performed excellently in detecting **sentiment** and tracking **topic evolution**, accurately reflecting major geopolitical shifts (2014, 2022).
- **Clustering Efficacy:** The combination of **Sentence-BERT, UMAP, and DBSCAN** emerged as a superior analytical tool. It allowed for a precise structural mapping of how the "enemy image" is constructed, which is often too nuanced for standard classification.
- **Reliability of Traditional Methods:** Basic **keyword frequency and regex analysis** remain highly effective and, in tasks requiring statistical precision, often outperform complex NLP models.

13 Team Contributions

- **Katarzyna Leniec** ($\approx$ 40h): EDA (Putin Corpus), Word frequency analysis, Propaganda detection, Reports.
- **Jan Pogłód** ($\approx$ 50h): EDA, preparation of a chapter on LLMs and RAG, comparison of different LLM models and their weaknesses. Preparation of a chapter on Putin's statements regarding Poland.
- **Cyprien Fourcroy** ($\approx$ 30h): Implementation of a RoBERTa-based sentiment analysis pipeline, sometimes combined with BERTopic for country- and topic-level sentiment analysis
- **Michał Puścian** ($\approx$ 30h): Development, analysis and presentation of BERTopic (chapter 7). Creating dashboard which allows everyone to explore topics at michalpuscian.com

References

Clayton Thomas Besaw . 2020. Unpacking Russian Presidential Speech Patterns with Machine Learning.

U. Abdullah and D. Alahmadi. 2022. Detecting propaganda techniques in English news articles using pretrained Transformers. *EMNLP workshop paper*.

Iuliia Alieva, Ian Kloo, and Kathleen M. Carley. 2024. Analyzing Russia's propaganda tactics on Twitter using mixed methods network analysis and natural language processing: a case study of the 2022 invasion of Ukraine. *EPJ Data Science*, 13(1):42.

M. Grootendorst. 2022. *BERTopic: Embedding-based topic modeling with BERT*.

HCDH-Uni-Heidelberg. 2025. *Chronorhetorics Corpus*.

N. Levshina. 2022. *Putin Corpus*.

Abdurahman Maarouf, Dominik Bär, Dominique Geissler, and Stefan Feuerriegel. 2024. HQP: A Human-Annotated Dataset for Detecting Online Propaganda. In *Findings of the Association for Computational Linguistics: ACL 2024*, pages 6064–6089, Bangkok, Thailand. Association for Computational Linguistics.

G. Martino, A. Barrón-Cedeno, H. Wachsmuth, R. Petrov, and P. Nakov. 2020. SemEval-2020 Task 11: Detection of propaganda techniques in news articles.

M. Mendonça and A. Figueira. 2025. Modeling Political Discourse with SentenceBERT and BERTopic.

M. Mochtak, P. Rupnik, and N. Ljubešić. 2024. The ParlaSent Multilingual Training Dataset for Sentiment Identification in Parliamentary Proceedings. *LREC-COLING 2024*.

K. Sprenkamp, D. G. Jones, and L. Zavolokina. 2023. Large Language Models for Propaganda Detection.

Touché. 2025. *Ideology and Power Identification in Parliamentary Debates*. CLEF 2025.

T. Wolf, L. Debut, V. Sanh, et al. 2020. Transformers: State-of-the-Art Natural Language Processing. *EMNLP Demonstration*.