Project Proposal

SNA, 2023

Group: Dasha Gorokhova, Elina Kalinina, Anastasia Rezova

21/06/23

https://github.com/LnKlnn/Entity-Linking

Points of discourse identification using knowledge graphs in presidential speeches. The case of President Putin's addresses to the Federal Assembly and SPIEF in 2023.

Abstract

The political landscape of Russia has been a topic of interest and concern for many years, especially in light of the country's growing influence on the global stage. This project aims to analyze the content of the speeches given by a leader of Russia, in order to gain a deeper understanding of the underlying structure of Russian policy. By leveraging social network analysis methods, the project will construct networks using discourse graphs to identify key players and priorities within the government, as well as potential areas for cooperation or conflict with other countries. The data for the project will be extracted from the Presidential Address of Vladimir Putin to the Federal Assembly on February 21, 2023 and speech at the St. Petersburg International Economic Forum (SPIEF) on June 16, 2023. The expected results include the construction of entity communities using discourse-based knowledge graphs, and the comparison of empirical network features to identify important directions of Russian policy.

Introduction

The concept of a knowledge graph has revolutionized the way search engines understand and deliver search results. Knowledge graphs are essentially graphs that represent relationships between various entities, and can be used for a wide range of research applications. In this project, the authors will use knowledge graphs to analyze the Presidential Address of Vladimir Putin to the Federal Assembly on February 21, 2023 and speech of President Putin at the St. Petersburg International Economic Forum (SPIEF) on June 16, 2023. By constructing networks based on the text, the authors will identify entity communities and important directions of Russian policy. This project will serve as a foundation for potential further research in the field.

Literature Review

The words 'knowledge graph' are very often followed by another phrase: "things, not strings." This is how Google introduced the new working principles of their search engine in 2012: it would now base the search results based on rich understanding of what particular search queries might mean and deliver different results based on these assumptions. What defines this new model is called a knowledge graph — a "'graph' that understands

real-world entities and their relationships to one another" (Singhal, 2012). Such graphs are a useful tool for representing relationships between various concepts, people, events, organizations, and many other entities.

Knowledge graphs can be used practically at any stage of research, even as an end product, but more often they are employed as a stepping stone for developing new services and furthering the research. For example, they can be used for identifying ideological leanings of the speaker using modified, opinion-aware knowledge graphs (Chen et. al., 2017) , or for storing useful knowledge, for example, regarding parliamentary speeches and debates, which can be used for further research regarding main topics of discourse (Tamper et. al., 2022) or related to press coverage referring to minorities (Martins et. al., 2021). In general, building a knowledge graph is often a first step when attempting to disentangle the relationship between the entities in a particular 'network', and this project will attempt to lay the ground for potential further research.

SNA Methods

To complete this project we are going to collect and prepare the data for analysis, preprocess data, construct networks using discourse-based knowledge graphs and then analyze these networks.

The data for the project is extracted from the text of the Presidential Address of Vladimir Putin to the Federal Assembly on February 21, 2023 and speech of President Putin at the St. Petersburg International Economic Forum (SPIEF) on June 16, 2023 . We are going to use the official English versions of President Putin's speeches received from the official site of the Kremlin (http://en.kremlin.ru/events/president/news/70565 http://en.kremlin.ru/events/president/news/71445). Thus, we have two speeches: the first speech primarily describes foreign policy while the second one describes domestic politics. And together these two texts can give all necessary information about all main points of the modern policy of the Russian Federation.

To construct networks we are going to preprocess these texts by natural language processing technology: firstly, we will transform sentences into SOV form (subject-object-verb). To do it we copied the speeches to the file input_search.txt and opened it in a Python program, then read the text. After that we divided it into sentences and used the NLP library (stanza) to transform each sentence into the SOV form. That is the temporary version of the text containing countries, geographical names, persons that play an important role in the modern policy of the Russian Federation and relations between them. Based on this version of the text we will build a network: nodes will be countries, geographical names, international organizations and persons, connections between them will be edges. After cleaning the artifacts, we compiled data frames and draw the network.

Results

By applying the methods mentioned in the above section, we built a knowledge graph that describes the concepts that were present in two presidential speeches in 2023 and relationships between them. Figure 1 shows the whole structure of the network.

FIGURE 1

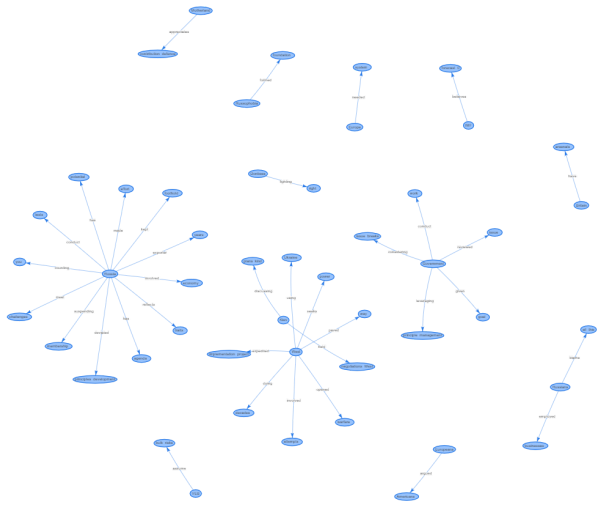

Table 1 shows the general information about the network.

TABLE 1

Number of nodes	49
Number of edges	36

Firstly, the dataset contains 23983 words and 144911 characters without spaces. The network is relatively small, with less than 50 nodes, and it is quite sparse, the network is directed. This, however, allows us to easily identify key concepts that are used in those two speeches.

The first of the influential concepts is that of ‘Russia’ and ‘Government,’ unsurprisingly. It is connected to (and therefore described by) the concepts like ‘potential,’ ‘challenges,’ which fall into phrases ‘Russia has potential’ and ‘Russia meets challenges.’ Other big groups of phrases which fall into a pattern are those relating to the United States and Europe. The phrases seem to be filled with a more charged rhetoric and emotions which allows us to assume with more conviction the sentiment which lies behind those words. These are phrases such as ‘West using Ukraine,’ ‘West opened warfare,’ and ‘West seeks power.’

From the word's frequency we can see that the words Russia and Western are used an equal number of times (16 times), which together with the visualisation on the graph suggest the the rhetoric of the discourse of the presidential speeches of 2023 contributes roughly the same attention both the state of Russia itself and to putting it in opposition to the outside world.

Conclusion

In today's world, where the amount of data is growing exponentially, it has become increasingly challenging to make sense of it all. Knowledge graphs have emerged as a powerful tool for organizing and analyzing complex information. A knowledge graph is a type of database that captures the relationships between entities and concepts, allowing for more efficient and effective analysis of data.

In recent years, knowledge graphs have been used in a variety of fields, including healthcare, finance, and government. One area where knowledge graphs have shown particular promise is in the analysis of political speeches. By creating a knowledge graph of a politician's speeches, researchers can identify the key themes and priorities of their policies.

This project demonstrates the potential of using knowledge graphs to analyze presidential speeches. Specifically, we analyzed the speeches of Vladimir Putin in 2023 to identify the key directions of Russian policy. Our analysis revealed the importance of Russia's relationship with the West, as well as its focus on economic growth and technological innovation.

The use of knowledge graphs in this project allowed us to identify patterns and connections that might have been missed using traditional methods of analysis. By visualizing the relationships between concepts and entities, we were able to gain a deeper understanding of Putin's priorities and goals.

The findings from this project serve as a foundation for further research in this field. The use of knowledge graphs in political analysis has the potential to revolutionize the way we understand and interpret political discourse. By identifying key themes and priorities, researchers can better understand the motivations and goals of political leaders.

Overall, this project demonstrates the power and potential of knowledge graphs in a variety of research applications. As the amount of data continues to grow, knowledge graphs will become an increasingly important tool for making sense of complex information.

References:

Chen, W., Zhang, X., Wang, T., Yang, B., & Li, Y. (2017, August). Opinion-aware Knowledge Graph for Political Ideology Detection. In *IJCAI* (Vol. 17, pp. 3647-3653).

Martins, P., Costa, L., & Ramalho, J. C. (2021). Knowledge graph of press clippings referring social minorities.

Singhal, Amit. (2012, May). Introducing the Knowledge Graph: things, not strings.

Tamper, M., Leal, R., Sinikallio, L., Leskinen, P., Tuominen, J., & Hyvönen, E. (2022, August). Extracting knowledge from parliamentary debates for studying political culture and language. In *Proceedings of the 1st International Workshop on Knowledge Graph Generation From Text and the 1st International Workshop on Modular Knowledge (TEXT2KG 2022 and MK2022)*. CEUR-WS. org.