Discuss Data

Open Platform for the Interactive Discussion of Research Data Quality
(on the example of area studies on the post-Soviet region)

The President's Words: A Term Frequency Analysis of Putin's and Medvedev's Statements in Official Kremlin Transcripts

Version 1.1, Dataset created by

https://discuss-data.net/dataset/7e016752-2438-46c9-9526-d35d41c823a2/

Data Citation

dekoder.org (2023): The President's Words: A Term Frequency Analysis of Putin's and Medvedev's Statements in Official Kremlin Transcripts, v. 1.1, Discuss Data, <doi: 10.48320/7E016752-2438-46C9-9526-D35D41C823A2>.

Description

This dataset consists of two parts: (1) A database of more than 10,000 transcripts published on the the official website of the President of Russia since 1999, December 31st (last update 2023, January 16th). (2) A frequency analysis of the terms used by president Putin and president Medvedev in these documents. This is the data used in the word counting tool at https://putin.dekoder.org/. New in version 1.1: This version has been updated to include texts up to 16 January 2023. Furthermore, it now also includes filtered versions of the original transcripts (with only Putin's and Medvedev's words left) and pre-processed wordlists (lemmatised versions of the filtered transcripts) as used in the analysis.

Files

Documentation of Data Collection.pdf (data documentation)

terms_en_1999-2023.csv.zip (data)

terms_ru_1999-2023.csv.zip (data)

transcripts_en_2023-01-16.json.zip (data)

transcripts_ru_2023-01-16.json.zip (data)

Metadata

Title:	The President's Words: A Term Frequency Analysis of Putin's and Medvedev's Statements in Official Kremlin Transcripts
Subtitle:	
Version:	1.1
Creators:	dekoder.org
Uploaded by:	Daniel Marcus
Main Category:	Elites
Additional Categories:	
Institutional Affiliation:	dekoder.org and Research Centre for East European Studies at the University of Bremen
Publication date:	Jan. 17, 2023, 11:45 a.m.
Period of data creation/gathering:	from March 1, 2020 to Jan. 16, 2023
Date of data creation (text):	Created in March 2020, last updated in January 2023.
Time period covered:	from Dec. 31, 1999 to Jan. 16, 2023
Time period covered (text):	
Sources of data:	The official Kremlin website: - Russian: http://kremlin.ru/events/ president/transcripts - English http://en.kremlin.ru/events/ president/transcripts
License:	Attribution: Open Data Commons Attribution License (ODC-By) v1.0
Data types:	statistics, text document
Data type (text):	
Countries:	Russia
Languages:	Russian, English
Disciplines:	Political Science
Keywords:	Kremlin, Medvedev, Putin, Russia
Related datasets:	
Related datasets (text):	

Related publications:

Dmitry Kartsev, Medvedev's Highlights, 2020, https://putin.dekoder.org/putin-medvedev/

Eduard Klein, The Fight Against Corruption: Much Ado About Nothing, 2020, https://putin.dekoder.org/corruption

Ekaterina Paustyan, Das Verhältnis zwischen dem Zentrum und den Regionen in der Rhetorik des russischen Präsidenten und der Staatsduma, in: Russland-Analysen Nr. 413, 2022, https://www.laender-analysen.de/russland-analysen/413/das-verhaeltnis-zwischen-dem-zentrum-und-den-regionen-in-der-rhetorik-des-russischen-praesidenten-und-der-staatsduma/

Gwendolyn Sasse, Die Ukraine in der Rhetorik russischer Präsidenten und der Staatsduma, in: Russland-Analysen Nr. 413, 2022, https://www.laender-analysen.de/russland-analysen/413/die-ukraine-in-der-rhetorik-russischer-praesidenten-und-der-staatsduma/

Gwendolyn Sasse, Russia and Ukraine, 2020, https://putin.dekoder.org/russia-ukraine

Gwendolyn Sasse, Russia and Ukraine in Kremlin Rhetoric, in: Russian Analytical Digest No. 250, 2000, https://css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/RAD250.pdf

Heiko Pleines, Die sozialen Sorgen der Bevölkerung in der politischen Rhetorik, in: Russland-Analysen Nr. 413, 2022, https://www.laender-analysen.de/russland-analysen/413/die-sozialen-sorgen-der-bevoelkerung-in-der-politischen-rhetorik/

Heiko Pleines, Social policy: Family above all?, 2020, https://putin.dekoder.org/social-policy

Ulrich Schmid, Nationalitätenpolitik: Russländische Nation versus russisches Volk?, in: Russland-Analysen Nr. 413, 2022, https://www.laender-analysen.de/russland-analysen/413/nationalitaetenpolitik-russlaendische-nation-versus-russisches-volk/

Ulrich Schmid, Nationality Policy: Russian Nation vs. Russian People?, in: Russian

Related projects:	This data collection was compiled within the project "Knowledge transfer squared - Russian studies" at the Research Centre for East European Studies at the University of Bremen.
Institutional affiliation:	dekoder.org and Research Centre for East European Studies at the University of Bremen
Funding:	The compilation of this data was funded by the Volkswagen Foundation.
Methods of data collection:	Web Scraping
Methods of data analysis:	Quantitative Content Analysis, Raw Data For Analysis

Version History

The President's Words: A Term Frequency Analysis of Putin's and Medvedev's Statements in Official Kremlin Transcripts, v1.0

Publications

Dmitry Kartsev, Medvedev's Highlights, 2020, https://putin.dekoder.org/putin-medvedev/, https://putin.dekoder.org/putin-medvedev/ (uses dataset)

Gwendolyn Sasse, Russia and Ukraine, 2020, https://putin.dekoder.org/russia-ukraine, https://putin.dekoder.org/russia-ukraine (uses dataset)

Heiko Pleines, Social policy: Family above all?, 2020, https://putin.dekoder.org/social-policy, https://putin.dekoder.org/social-policy (uses dataset)

Eduard Klein, The Fight Against Corruption: Much Ado About Nothing, 2020, https://putin.dekoder.org/corruption, https://putin.dekoder.org/corruption (uses dataset)

Heiko Pleines, Die sozialen Sorgen der Bevölkerung in der politischen Rhetorik, in: Russland-Analysen Nr. 413, 2022, https://www.laender-analysen.de/russland-analysen/413/die-sozialen-sorgen-der-bevoelkerung-in-der-politischen-rhetorik/, https://www.laender-analysen.de/russland-analysen/413/die-sozialen-sorgen-der-bevoelkerung-in-der-politischen-rhetorik/ (uses dataset)

Ekaterina Paustyan, Das Verhältnis zwischen dem Zentrum und den Regionen in der Rhetorik des russischen Präsidenten und der Staatsduma, in: Russland-Analysen Nr. 413, 2022, https://www.laender-analysen.de/russland-analysen/413/das-verhaeltnis-zwischen-dem-zentrum-und-den-regionen-in-der-rhetorik-des-russischen-praesidenten-und-der-staatsduma/, https://www.laender-analysen.de/russland-analysen/413/das-verhaeltnis-zwischen-dem-zentrum-und-den-regionen-in-der-rhetorik-des-russischen-praesidenten-und-der-staatsduma/ (uses dataset)

Ulrich Schmid, Nationalitätenpolitik: Russländische Nation versus russisches Volk?, in: Russland-Analysen Nr. 413, 2022, https://www.laender-analysen.de/russland-analysen/413/nationalitaetenpolitik-russlaendische-nation-versus-russisches-volk/, https://www.laender-analysen.de/russland-analysen/413/nationalitaetenpolitik-russlaendische-nation-versus-russisches-volk/ (uses dataset)

Yana Zabanova, Russland entdeckt die Energiewende: Ein Sonderweg zur Dekarbonisierung?, in: Russland-Analysen Nr. 413, 2022, https://www.laender-analysen.de/russland-analysen/413/russland-entdeckt-die-energiewende-ein-sonderweg-zur-dekarbonisierung/, https://www.laender-analysen.de/russland-analysen/413/russland-entdeckt-die-energiewende-ein-sonderweg-zur-dekarbonisierung/ (uses dataset)

Gwendolyn Sasse, Die Ukraine in der Rhetorik russischer Präsidenten und der Staatsduma, in: Russland-Analysen Nr. 413, 2022, https://www.laender-analysen.de/russland-analysen/413/die-ukraine-in-der-rhetorik-russischer-praesidenten-und-der-staatsduma/, https://www.laender-analysen.de/russland-analysen/413/die-ukraine-in-der-rhetorik-russischer-praesidenten-und-der-staatsduma/ (uses dataset)

Gwendolyn Sasse, Russia and Ukraine in Kremlin Rhetoric, in: Russian Analytical Digest No. 250, 2000, https://css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/RAD250.pdf, https://css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/RAD250.pdf (uses dataset)

Ulrich Schmid, Nationality Policy: Russian Nation vs. Russian People?, in: Russian Analytical Digest No. 250, 2000, https://css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/RAD250.pdf, https://css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/RAD250.pdf (uses dataset)

Data License

Attribution: Open Data Commons Attribution License (ODC-By) v1.0 (standard)

Data Collections with the Open Data Commons Attribution License (ODC-BY) can be freely

- **assessed and used**
- **reused**
- **and redistributed**

if **proper attribution** is provided.

Any public use of the database, or works produced from the database, must be attributed in the manner specified in the license. For any use or redistribution of the database, or works produced from it, the license of the database must be clear and any notices on the original database must be kept intact.

For more information, please read the Open Data Commons Attribution License (ODC-BY) Full Legal Text.