The UN Security Council debates on Afghanistan 2001-2017

Mirco Schönfeld, Steffen Eckhard, Ronny Patz, Hilde van Meegdenburg

June 2018

1 Description

This dataset contains meeting protocols from the United Nations Security Council meetings on "The situation in Afghanistan" between 2001–2017.

txt.tar.gz This archive contains text file representations of the official meeting protocols.

speeches.tar.gz This is an archive of text files with single speeches in them. Every text file's filename contains an identifier for the meeting, a year, and an ordering number for every speech of the corresponding meeting. The first part of the filename equals the filename of the corresponding meeting in `txt.tar.gz`.

UNSC_Afghan_SpchCntr.tsv This is a tabulator-separated text file containing metadata for every speech. See section 4 for an explanation of fields.

UNSC_Afghan_Spchs_Meta.RData This is an `R`-object containing metadata as well as texts of all speeches. See section 4 for an explanation of fields. The object's structure allows for an easy conversion to a `tm`-Corpus object for further processing, for example.

2 Citation

When using our dataset, please use the citation indicated on the Dataverse's website.

> Schoenfeld, Mirco; Eckhard, Steffen; Patz, Ronny; Meegdenburg, Hilde van, 2018, "The UN Security Council debates on Afghanistan", `https://doi.org/10.7910/DVN/OM9RG8`, Harvard Dataverse, V1

3 Sources

This dataset is compiled semi-automatically from meeting protocols downloaded from `http://www.un.org/en/sc/meetings/`

Please contact mirco.schoenfeld@tum.de for any inaccuracies you may encounter.

4 Explanation of variables

year – The year of the speech

spv – The "S/PV"-meeting record identifier. The correct identifier is constructed by S/PV.{entry}. If the entry contains 'Resumption', the correct identifier is constructed by encapsulating the 'Resumption'-part as well as the following number in parentheses.

speech – An ongoing number numbering the speeches of one meeting.

speaker – The name of the speaker in lower case letters.

country – The country or affiliation the speaker represents during the speech. This is indicated either in the participants list of a meeting protocol or the beginning of the speech if the speaker is not listed as part of the participants list.

role – Expresses the role of the speaker. This field has one of the following three entries:

- `president`: The speaker was the meeting's president
- `listed`: The speaker is listed in the participants list
- `guest`: The speaker is not listed in the participants list but rather invited to take a seat at the table during the meeting.

text – **only in R dataset** – The actual text of the speech

doc – The filename of the speech given in `speeches.tar.gz`.

meetingdoc – The filename of the meeting protocol in `txt.tar.gz`.