

# Package ‘sotu’

May 9, 2026

**Type** Package

**Title** United States Presidential State of the Union Addresses

**Version** 1.0.4

**Author** Taylor B. Arnold [aut, cre]

**Maintainer** Taylor B. Arnold <tarnold2@richmond.edu>

**Description** The President of the United States is constitutionally obligated to provide a report known as the 'State of the Union'. The report summarizes the current challenges facing the country and the president's upcoming legislative agenda. While historically the State of the Union was often a written document, in recent decades it has always taken the form of an oral address to a joint session of the United States Congress. This package provides the raw text from every such address with the intention of being used for meaningful examples of text analysis in R. The corpus is well suited to the task as it is historically important, includes material intended to be read and material intended to be spoken, and it falls in the public domain. As the corpus spans over two centuries it is also a good test of how well various methods hold up to the idiosyncrasies of historical texts. Associated data about each address, such as the year, president, party, and format, are also included.

**Depends** R (>= 3.5.0)

**Imports** utils

**License** GPL-2

**URL** <https://github.com/statsmaths/sotu/>

**BugReports** <https://github.com/statsmaths/sotu/issues/>

**LazyData** true

**RoxygenNote** 7.2.1

**NeedsCompilation** no

**Repository** CRAN

**Date/Publication** 2022-08-17 13:40:02 UTC

## Contents

sotu-package . . . . .	2
sotu_dir . . . . .	3
sotu_meta . . . . .	3
sotu_text . . . . .	4
<b>Index</b>	<b>5</b>

---

sotu-package	<i>sotu: State of the Union Addresses</i>
--------------	---

---

## Description

The President of the United States is constitutionally obligated to provide a report of the current state of country. While historically this could be a written document, in modern times this has always taken the form of an oral address to a joint session of the US Congress. This package provides the raw text from every State of the Union address, with the intention of being used for meaningful examples of text analysis in R. This corpus is well-suited to the task as it is historically important, includes both material intended to be read and material intended to be spoken, and it falls in the public domain. As the corpus spans over two centuries it is also a good test of how well various methods hold up to the idiosyncrasies of historical texts.

## Details

A metadata table containing information about each speech is also included in the package. This includes information about the date and president giving the address. Note that some of the addresses contained in the archive are not technically State of the Union Addresses, but are more general Addresses to a Joint Session of Congress. Usually these occur following a new president's inauguration, and are historically treated as no different than a standard State of the Union.

## Author(s)

**Maintainer:** Taylor B. Arnold <tarnold2@richmond.edu>

## See Also

Useful links:

- <https://github.com/statsmaths/sotu/>
- Report bugs at <https://github.com/statsmaths/sotu/issues/>

---

`sotu_dir`*Save State of the Union Addresses in Directory*

---

**Description**

This function saves the State of the Union Addresses to disk, with one address in each file. Useful for testing and demonstrating text analysis functions that read from files.

**Usage**

```
sotu_dir(dir = tempfile())
```

**Arguments**

`dir` an optional directory of where the files should be saved. Otherwise, a temporary location is used. Will be created if it does not exist. Note that these temporary files are erased (eventually) after exiting R.

**Value**

a character vector with complete paths to the texts

**Author(s)**

Taylor B. Arnold, <taylor.arnold@acm.org>

**Examples**

```
fp <- sotu_dir()
txt <- readLines(fp[200])
txt
```

---

`sotu_meta`*Metadata from State of the Union Addresses*

---

**Description**

A data frame containing one row per State of the Union Address. These line up with the elements in [sotu\\_text](#).

**References**

The American Presidency Project, UCSB

---

`sotu_text`*State of the Union Address Text*

---

**Description**

A character vector with one State of the Union in each element. These line up with the rows of the data in [sotu\\_meta](#).

**References**

The American Presidency Project, UCSB

# Index

## \* data

sotu\_meta, 3

sotu\_text, 4

sotu (sotu-package), 2

sotu-package, 2

sotu\_dir, 3

sotu\_meta, 3, 4

sotu\_text, 3, 4