

# Package ‘sourcetools’

March 28, 2026

**Type** Package

**Title** Tools for Reading, Tokenizing and Parsing R Code

**Version** 0.1.7-2

**Maintainer** Kevin Ushey <kevinushey@gmail.com>

**Description** Tools for the reading and tokenization of R code. The 'sourcetools' package provides both an R and C++ interface for the tokenization of R code, and helpers for interacting with the tokenized representation of R code.

**License** MIT + file LICENSE

**Depends** R (>= 3.0.2)

**Suggests** testthat

**RoxygenNote** 5.0.1

**BugReports** <https://github.com/kevinushey/sourcetools/issues>

**Encoding** UTF-8

**NeedsCompilation** yes

**Author** Kevin Ushey [aut, cre]

**Repository** CRAN

**Date/Publication** 2026-03-28 06:10:35 UTC

## Contents

read . . . . .	2
tokenize_file . . . . .	2
<b>Index</b>	<b>4</b>

---

read	<i>Read the Contents of a File</i>
------	------------------------------------

---

**Description**

Read the contents of a file into a string (or, in the case of `read_lines`, a vector of strings).

**Usage**

```
read(path)
read_lines(path)
read_bytes(path)
read_lines_bytes(path)
```

**Arguments**

path	A file path.
------	--------------

---

tokenize_file	<i>Tokenize R Code</i>
---------------	------------------------

---

**Description**

Tools for tokenizing R code.

**Usage**

```
tokenize_file(path)
tokenize_string(string)
tokenize(file = "", text = NULL)
```

**Arguments**

file, path	A file path.
text, string	R code as a character vector of length one.

**Value**

A data.frame with the following columns:

value	The token's contents, as a string.
row	The row where the token is located.
column	The column where the token is located.
type	The token type, as a string.

**Note**

Line numbers are determined by existence of the `\n` line feed character, under the assumption that code being tokenized will use either `\n` to indicate newlines (as on modern Unix systems), or `\r\n` as on Windows.

**Examples**

```
tokenize_string("x <- 1 + 2")
```

# Index

`read`, [2](#)  
`read_bytes (read)`, [2](#)  
`read_lines (read)`, [2](#)  
`read_lines_bytes (read)`, [2](#)  
  
`tokenize (tokenize_file)`, [2](#)  
`tokenize_file`, [2](#)  
`tokenize_string (tokenize_file)`, [2](#)