

# Package ‘rbibutils’

July 21, 2020

**Type** Package

**Title** Convert Between Bibliography Formats

**Version** 1.0.3

**Date** 2020-07-20

**Description** Converts between a number of bibliography formats, including 'BibTeX', 'BibLaTeX' and 'Bibentry'. Includes a port of the 'bibutils' utilities by Chris Putnam <<https://sourceforge.net/projects/bibutils/>>. Supports all bibliography formats and character encodings implemented in 'bibutils'.

**License** GPL-2

**URL** <https://geobosh.github.io/rbibutils/> (website),  
<https://github.com/GeoBosh/rbibutils> (devel)

**BugReports** <https://github.com/GeoBosh/rbibutils/issues>

**Depends** R (>= 2.10)

**Imports** utils, tools, xml2

**Encoding** UTF-8

**LazyData** true

**NeedsCompilation** yes

**Suggests** testthat

**Author** Georgi N. Boshnakov [aut, cre] (R port, modifications to bibutils' C code),  
Chris Putman [aut] (src/\*, author of the bibutils libraries, <https://sourceforge.net/projects/bibutils/>),  
Richard Mathar [ctb] (src/addsout.c),  
Johannes Wilm [ctb] (src/biblatexin.c, src/bltypes.c)

**Maintainer** Georgi N. Boshnakov <[georgi.boshnakov@manchester.ac.uk](mailto:georgi.boshnakov@manchester.ac.uk)>

**Repository** CRAN

**Date/Publication** 2020-07-21 00:00:03 UTC

## R topics documented:

|                             |   |
|-----------------------------|---|
| rbibutils-package . . . . . | 2 |
| bibConvert . . . . .        | 3 |

|              |          |
|--------------|----------|
| <b>Index</b> | <b>7</b> |
|--------------|----------|

rbibutils-package      *Convert Between Bibliography Formats*

### Description

Converts between a number of bibliography formats, including 'BibTeX', 'BibLaTeX' and 'Bibentry'. Includes a port of the 'bibutils' utilities by Chris Putnam <<https://sourceforge.net/projects/bibutils/>>. Supports all bibliography formats and character encodings implemented in 'bibutils'.

### Details

Package **rbibutils** provides an R port of the bibutils programs plus additional facilities. The main function, `bibConvert`, offers all conversions between bibliography formats supported by library `bibutils`. In addition, package **rbibutils** converts to and from R's `bibentry` Bibtex-based bibliography format.

The core functionality is provided by the `bibutils` programs which convert between various bibliography formats using a common MODS XML intermediate format, see the source cited below.

Currently we provide the function `bibConvert` for conversion between supported bibliography formats. For complete list of formats supported by the package, see the documentation of the original `bibutils` library.

All encodings supported by the `bibutils` library are available for `bibConvert`.

Further functionality may be provided in future releases, in particular, the underlying C functions could be exposed to package authors. Further R wrappers may be added, as well. However, the scope of the package will remain conversion between formats based on `bibutils` and manipulation of the MODS XML intermediate format. **rbibutils** can be used also as an alternative to package **bibtex** (Francois 2020). For bibliography management see package **RefManager** (McLean 2017). For citations in R documentation (Rd or roxygen2) see package **Rdpack** (Boshnakov 2020).

The only external dependency (import) is package **xml2** and conversion between formats not involving `bibentry` do not need it, so it could be made "suggested" if there is a strong use case for that.

### Note

The `bibutils` library is included in a number of software packages. These include include `pandoc` and a library for Haskell. Executable programs for conversion are available for Linux distributions but seem not easily available for Windows. Executable and libraries can also be generated out-of-the-box from the `bibutils` distribution (on Windows under MSYS).

### Author(s)

Georgi N. Boshnakov (R code and R port of `bibutils`), Chris Putnam (author of `bibutils` library)

## References

Boshnakov GN (2020). “Rdpack: Update and Manipulate Rd Documentation Objects.” doi: [10.5281/zenodo.3925612](https://doi.org/10.5281/zenodo.3925612) (URL: <https://doi.org/10.5281/zenodo.3925612>), R package version 1.0.0.

Francois R. (2020). bibtex: Bibtex Parser. R package version 0.4.2.2. <https://CRAN.R-project.org/package=bibtex>.

McLean MW (2017). “RefManageR: Import and Manage BibTeX and BibLaTeX References in R.” *The Journal of Open Source Software*. doi: [10.21105/joss.00338](https://doi.org/10.21105/joss.00338) (URL: <https://doi.org/10.21105/joss.00338>).

Putnam C. (2003–2020), Library bibutils, <https://sourceforge.net/projects/bibutils/>.

## See Also

[bibConvert](#) for further details and examples

---

|            |   |
|------------|---|
| bibConvert | <i>Convert between bibliography formats</i> |
|------------|---|

---

## Description

Read a bibliography file in one of the supported formats, convert it to another format, and write it to a file.

## Usage

```
bibConvert(infile, outfile, informat, outformat, ..., tex, encoding,
           options)
```

## Arguments

|           |  |
|-----------|--|
| infile    | input file, a character string.  |
| outfile   | output file, a character string.   |
| informat  | input format, a character string, see Details.   |
| outformat | output format, a character string, see Details.  |
| ...       | not used.  |
| tex       | TeX specific options, see Details, a character vector.   |
| encoding  | character(2), a length two vector specifying input and output encodings. Default to both is "utf8", see Details. |
| options   | mainly for debugging: additional options for the converters, see Details.  |

## Details

Arguments `informat` and `outformat` can usually be omitted, since `bibConvert` infers them from the extensions of the names of the input and output files. However, there is ambiguity for the extension `"bib"`, since it is used for Bibtex and Biblatex entries. For this extension, the default for both, `informat` and `outformat`, is `"bibtex"`.

Package **`rbibutils`** supports format `"bibentry"`, in addition to the formats supported by the `bibutils` library. A `bibentry` object contains one or more references. Two formats are supported for `"bibentry"` for both input and output. A `bibentry` object previously saved to a file using `saveRDS` (default extension `"rds"`) or an R source file containing one or more `bibentry` commands. The `"rds"` file is just read in and should contain a `bibentry` object.

When `bibconvert` outputs to an R source file, there is one `bibentry` call for each reference, just as in a Bibtex file, each reference is a single entry. Such a file can be used as input to `bibConvert`. For input `bibConvert` accepts also R code containing additional instructions. The input file is (effectively) source'd, all `bibentry` objects created by it are collected and merged into a single `bibentry` object.

Note that when the input format and output format are identical, the conversion is not necessarily a null operation (except for `xml`, and even that may change). It should be remembered also that there may be loss of information when converting from one format to another.

For complete list of supported bibliography formats, see the documentation of the original `bibutils` library.

Argument `encoding` is a character vector containing 2 elements, specifying the encoding of the input and output files. If the encodings are the same, a length one vector can be supplied. The default encodings are UTF-8 for input and output. A large number of familiar encodings are supported, e.g. `"latin1"` and `"cp1251"` (Windows Cyrillic). Some encodings have two or more aliases and they are also accepted. If an unknown encoding is requested, a list of all supported encodings will be printed.

Argument `tex` is an unnamed character vector containing switches for bibtex input and output (mostly output). Currently, the following are available:

**`uppercase`** write bibtex tags/types in upper case.

**`no_latex`** do not convert latex-style character combinations to letters.

**`brackets`** use brackets, not quotation marks surrounding data.

**`dash`** use one dash `"-"`, not two `--"`, in page ranges.

**`fc`** add final comma to bibtex output.

By default latex encodings for accented characters are converted to letters. This may be a problem if the output encoding is not UTF-8, since some characters created by this process may be invalid in that encoding. For example, a Bibtex file which otherwise contain only cyrillic and latin characters may have a few entries with authors containing latin accented characters represented using the TeX convention. If those characters are not converted to Unicode letters, they can be exported to `"cp1251"` (Windows Cyrillic) for example. Specifying the option `no_latex` should solve the problem in such cases.

Argument `options` is mostly for debugging and mimics the command line options of the `bibutils`' binaries. The argument is a named character vector and is supplied as `c(tag1= val1, tag2= val2, ...)`, where each `tag` is the name of an option and the value is the corresponding value. The

value for options that do not require one is ignored and can be set to "". Some of the available options are:

**h** help, show all available options.

**nb** do not write Byte Order Mark in UTF8 output.

**verbose** print intermediate output.

**debug** print even more intermediate output.

### Value

The function is used for the side effect of creating a file in the requested format. It returns a list, currently containing the following components:

|                       |  |
|-----------------------|--|
| <code>infile</code>   | name of the input file,                          |
| <code>outfile</code>  | name of the output file,                         |
| <code>nref_in</code>  | number of references read from the input file,   |
| <code>nref_out</code> | number of references written to the output file. |

### Author(s)

Georgi N. Boshnakov

### References

Chris Putnam, Library bibutils, <https://sourceforge.net/projects/bibutils/>.

### Examples

```
fn_biblatex <- system.file("bib", "ex0.biblatex", package = "rbibutils")
fn_biblatex
## file.show(fn_biblatex)

## convert a biblatex file to xml
mod1 <- tempfile(fileext = ".xml")
bibConvert(infile = fn_biblatex, outfile = mod1, informat = "biblatex", outformat = "xml")
## file.show(mod1)

## convert a biblatex file to bibtex
bib <- tempfile(fileext = ".bib")
bib2 <- tempfile(fileext = ".bib")
bibConvert(infile = fn_biblatex, outfile = bib, informat = "biblatex", outformat = "bib")
## file.show(bib)

## convert a biblatex file to bibentry
rds <- tempfile(fileext = ".rds")
fn_biblatex
rds
be <- bibConvert(fn_biblatex, rds, "biblatex", "bibentry")
bea <- bibConvert(fn_biblatex, rds, "biblatex") # same
readRDS(rds)
```

```
## convert to R source file
r <- tempfile(fileext = ".R")
bibConvert(fn_biblatex, r, "biblatex")
file.show(r)
cat(readLines(r), sep = "\n")

fn_cyr_utf8 <- system.file("bib", "cyr_utf8.bib", package = "rbibutils")

## Can't have files with different encodings in the package, so below
## first convert a UTF-8 file to something else.
##
## input here contains cyrillic (UTF-8) output to Windows Cyrillic,
## notice the "no_latex" option
a <- bibConvert(fn_cyr_utf8, bib, encoding = c("utf8", "cp1251"), tex = "no_latex")

## now take the bib file and convert it to UTF-8
bibConvert(bib, bib2, encoding = c("cp1251", "utf8"))

## Latin-1 example: Author and Title fields contain Latin-1 accented
## characters, not real names. As above, the file is in UTF-8
fn_latin1_utf8 <- system.file("bib", "latin1accents_utf8.bib", package = "rbibutils")
## convert to Latin-1, by default the accents are converted to TeX combinations:
b <- bibConvert(fn_latin1_utf8, bib, encoding = c("utf8", "latin1"))
cat(readLines(bib), sep = "\n")
## use "no_latex" option to keep them Latin1:
c <- bibConvert(fn_latin1_utf8, bib, encoding = c("utf8", "latin1"), tex = "no_latex")
## this will show properly in Latin-1 locale (or suitable text editor):
##cat(readLines(bib), sep = "\n")

unlink(c(mod1, bib, bib2, r, rds))
```

# Index

\* **documentation**

    bibConvert, [3](#)

\* **package**

    rbibutils-package, [2](#)

bibConvert, [2](#), [3](#), [3](#)

rbibutils (rbibutils-package), [2](#)

rbibutils-package, [2](#)