

# Package ‘pkgcache’

April 18, 2020

**Title** Cache 'CRAN'-Like Metadata and R Packages

**Version** 1.0.7

**Description** Metadata and package cache for CRAN-like repositories.  
This is a utility package to be used by package management tools that want to take advantage of caching.

**License** MIT + file LICENSE

**Encoding** UTF-8

**LazyData** true

**URL** <https://github.com/r-lib/pkgcache#readme>

**BugReports** <https://github.com/r-lib/pkgcache/issues>

**Imports** assertthat, callr (>= 2.0.4.9000), cli (>= 2.0.0), curl (>= 3.2), digest, filelock, glue, prettyunits, R6, processx (>= 3.3.0.9001), rappdirs, rlang, tibble, tools, utils, uuid

**Suggests** covr, debugme, desc, fs, jsonlite, mockery, pingr, rprojroot, testthat, withr

**Depends** R (>= 3.1)

**RoxygenNote** 7.0.2

**NeedsCompilation** no

**Author** Gábor Csárdi [aut, cre]

**Maintainer** Gábor Csárdi <csardi.gabor@gmail.com>

**Repository** CRAN

**Date/Publication** 2020-04-18 05:20:08 UTC

## R topics documented:

pkgcache-package . . . . .	2
cranlike_metadata_cache . . . . .	5
get_cranlike_metadata_cache . . . . .	8
meta_cache_deps . . . . .	8
package_cache . . . . .	10
pkg_cache_summary . . . . .	12

---

pkgcache-package      *Cache for package data and metadata*

---

## Description

Metadata and package cache for CRAN-like repositories. This is a utility package to be used by package management tools that want to take advantage of caching.

## Details

Metadata and package cache for CRAN-like repositories. This is a utility package to be used by package management tools that want to take advantage of caching.

### Installation:

You can install the released version of pkgcache from [CRAN](#) with:

```
install.packages("pkgcache")
```

### Metadata cache:

`meta_cache_list()` lists all packages in the metadata cache. It includes Bioconductor package, and all versions (i.e. both binary and source) of the packages for the current platform and R version.

```
library(pkgcache)
meta_cache_list()
#>
#> i Loading global cached package metadata
#> v Loading global cached package metadata
#> # A tibble: 35,372 x 33
#>   package title version depends suggests built imports archs repodir
#>   <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr> <chr>
#> 1 A3      "Acc~ 1.0.0 R (>= ~ randomF~ R 3.~ <NA> <NA> bin/ma~
#> 2 aaSEA   Amin~ 1.1.0 R(>= 3~ knitr, ~ R 3.~ "DT(>=~ <NA> bin/ma~
#> 3 ABACUS  "App~ 1.0.0 R (>= ~ rmarkdo~ R 3.~ ggplot~ <NA> bin/ma~
#> 4 abbyyR  Acce~ 0.5.5 R (>= ~ testtha~ R 3.~ httr, ~ <NA> bin/ma~
#> 5 abc.da~ Data~ 1.0 R (>= ~ <NA> R 3.~ <NA> <NA> bin/ma~
#> 6 ABC.RAP Arra~ 0.9.0 R (>= ~ knitr, ~ R 3.~ graphi~ <NA> bin/ma~
#> 7 abc     Tool~ 2.1 R (>= ~ <NA> R 3.~ <NA> <NA> bin/ma~
#> 8 abcADM  Fit ~ 1.0 <NA> <NA> R 3.~ Rcpp (~ abcA~ bin/ma~
#> 9 ABCana~ Comp~ 1.2.1 R (>= ~ <NA> R 3.~ plotrix <NA> bin/ma~
#> 10 abcdeF~ "ABC~ 0.4 Rglpk,~ LIM,syb~ R 3.~ <NA> <NA> bin/ma~
#> # ... with 35,362 more rows, and 24 more variables: platform <chr>,
#> # rversion <chr>, needscompilation <chr>, priority <chr>, ref <chr>,
#> # type <chr>, direct <lgl>, status <chr>, target <chr>, mirror <chr>,
#> # sources <list>, filesize <int>, sha256 <chr>, sysreqs <chr>,
#> # published <dtm>, deps <list>, license <chr>, md5sum <chr>,
#> # linkingto <chr>, enhances <chr>, license_restricts_use <chr>,
#> # os_type <chr>, license_is_foss <chr>, path <chr>
```

`meta_cache_deps()` and `meta_cache_revdeps()` can be used to look up dependencies and reverse dependencies.

The metadata is updated automatically if it is older than seven days, and it can also be updated manually with `meta_cache_update()`.

See the `cranlike_metadata_cache` R6 class for a lower level API, and more control.

### Package cache:

Package management tools may use the `pkg_cache_*` functions and in particular the `package_cache` class, to make use of local caching of package files.

The `pkg_cache_*` API is high level, and uses a user level cache:

```
pkg_cache_summary()
#> $cachepath
#> [1] "/Users/gaborcsardi/Library/Caches/R-pkg/pkg"
#>
#> $files
#> [1] 783
#>
#> $size
#> [1] 724791904

pkg_cache_list()
#> # A tibble: 783 x 10
#>   fullpath path package url etag sha256 built version platform
#>   <chr> <chr> <chr> <chr> <chr> <chr> <int> <chr> <chr>
#> 1 /Users/~ src/~ crayon <NA> <NA> 84be6~ 0 <NA> <NA>
#> 2 /Users/~ bin/~ proces~ http~ "\"3~ ce84a~ NA 3.4.1 macos
#> 3 /Users/~ bin/~ R6 http~ "\"d~ 4cfe3~ NA 2.4.0 macos
#> 4 /Users/~ bin/~ ps http~ "\"3~ 63add~ NA 1.3.0 macos
#> 5 /Users/~ src/~ callr <NA> <NA> df85a~ 0 <NA> <NA>
#> 6 /Users/~ bin/~ callr http~ "\"5~ 1a5e3~ NA 3.3.1 macos
#> 7 /Users/~ bin/~ assert~ http~ "\"d~ 103f9~ NA 0.2.1 macos
#> 8 /Users/~ bin/~ backpo~ http~ "\"d~ 344bb~ NA 1.1.4 macos
#> 9 /Users/~ bin/~ crayon http~ "\"b~ 2ac4f~ NA 1.3.4 macos
#> 10 /Users/~ bin/~ desc http~ "\"4~ 196ce~ NA 1.2.0 macos
#> # ... with 773 more rows, and 1 more variable: vignettes <chr>

pkg_cache_find(package = "dplyr")
#> # A tibble: 5 x 10
#>   fullpath path package url etag sha256 built version platform
#>   <chr> <chr> <chr> <chr> <chr> <chr> <int> <chr> <chr>
#> 1 /Users/~ bin/~ dplyr http~ "\"6~ e2594~ NA 0.8.3 macos
#> 2 /Users/~ src/~ dplyr <NA> <NA> 134b1~ 0 <NA> <NA>
#> 3 /Users/~ src/~ dplyr <NA> <NA> b357f~ 0 <NA> <NA>
#> 4 /Users/~ src/~ dplyr <NA> <NA> 81767~ 0 <NA> <NA>
#> 5 /Users/~ src/~ dplyr <NA> <NA> b237c~ 0 <NA> <NA>
#> # ... with 1 more variable: vignettes <chr>
```

`pkg_cache_add_file()` can be used to add a file, `pkg_cache_delete_files()` to remove files, `pkg_cache_get_files()` to copy files out of the cache.

The `package_cache` class provides a finer API.

**Bioconductor support:**

Both the metadata cache and the package cache support Bioconductor by default, automatically. See the `BioC_mirror` option and the `R_BIOC_MIRROR` and `R_BIOC_VERSION` environment variables below to configure `pkgcache`'s Bioconductor support.

**Package Options:**

- The `BioC_mirror` option can be used to select a Bioconductor mirror. This takes priority over the `R_BIOC_MIRROR` environment variable.
- `pkgcache_timeout` is the HTTP timeout for all downloads. It is in seconds, and the limit for downloading the whole file. Defaults to 3600, one hour. It corresponds to the `TIMEOUT libcurl option`.
- `pkgcache_connecttimeout` is the HTTP timeout for the connection phase. It is in seconds and defaults to 30 seconds. It corresponds to the `CONNECTTIMEOUT libcurl option`.
- `pkgcache_low_speed_limit` and `pkgcache_low_speed_time` are used for a more sensible HTTP timeout. If the download speed is less than `pkgcache_low_speed_limit` bytes per second for at least `pkgcache_low_speed_time` seconds, the download errors. They correspond to the `LOW_SPEED_LIMIT` and `LOW_SPEED_TIME` curl options.

**Package environment variables:**

- The `R_BIOC_VERSION` environment variable can be used to override the default Bioconductor version detection and force a given version. E.g. this can be used to force the development version of Bioconductor.
- The `R_BIOC_MIRROR` environment variable can be used to select a Bioconductor mirror. The `BioC_mirror` option takes priority over this, if set.
- `PKGCACHE_TIMEOUT` is the HTTP timeout for all downloads. It is in seconds, and the limit for downloading the whole file. Defaults to 3600, one hour. It corresponds to the `TIMEOUT libcurl option`. The `pkgcache_timeout` option has priority over this, if set.
- `PKGCACHE_CONNECTTIMEOUT` is the HTTP timeout for the connection phase. It is in seconds and defaults to 30 seconds. It corresponds to the `CONNECTTIMEOUT libcurl option`. The `pkgcache_connecttimeout` option takes precedence over this, if set.
- `PKGCACHE_LOW_SPEED_LIMIT` and `PKGCACHE_LOW_SPEED_TIME` are used for a more sensible HTTP timeout. If the download speed is less than `PKGCACHE_LOW_SPEED_LIMIT` bytes per second for at least `PKGCACHE_LOW_SPEED_TIME` seconds, the download errors. They correspond to the `LOW_SPEED_LIMIT` and `LOW_SPEED_TIME` curl options. The `pkgcache_low_speed_time` and `pkgcache_low_speed_limit` options have priority over these environment variables, if they are set.

**Code of Conduct:**

Please note that the 'pkgcache' project is released with a [Contributor Code of Conduct](#). By contributing to this project, you agree to abide by its terms.

**License:**

MIT © [RStudio Inc](#)

**Author(s)**

**Maintainer:** Gábor Csárdi <[csardi.gabor@gmail.com](mailto:csardi.gabor@gmail.com)>

**See Also**

Useful links:

- <https://github.com/r-lib/pkgcache#readme>
- Report bugs at <https://github.com/r-lib/pkgcache/issues>

---

cranlike\_metadata\_cache

*Metadata cache for a CRAN-like repository*

---

**Description**

This is an R6 class that implements the metadata cache of a CRAN-like repository. For a higher level interface, see the [meta\\_cache\\_list\(\)](#), [meta\\_cache\\_deps\(\)](#), [meta\\_cache\\_revdeps\(\)](#) and [meta\\_cache\\_update\(\)](#) functions.

**Details**

The cache has several layers:

- The data is stored inside the `cranlike_metadata_cache` object.
- It is also stored as an RDS file, in the session temporary directory. This ensures that the same data is used for all queries of a `cranlike_metadata_cache` object.
- It is stored in an RDS file in the user's cache directory.
- The downloaded raw `PACKAGES*` files are cached, together with HTTP ETags, to minimize downloads.

It has a synchronous and an asynchronous API.

**Usage**

```
cmc <- cranlike_metadata_cache$new(  
  primary_path = NULL, replica_path = tempfile(),  
  platforms = default_platforms(), r_version = current_r_version(),  
  bioc = TRUE, cran_mirror = default_cran_mirror(),  
  repos = getOption("repos"),  
  update_after = as.difftime(7, units = "days"))  
  
cmc$list/packages = NULL  
cmc$async_list/packages = NULL  
  
cmc$deps/packages, dependencies = NA, recursive = TRUE  
cmc$async_deps/packages, dependencies = NA, recursive = TRUE  
  
cmc$revdeps/packages, dependencies = NA, recursive = TRUE  
cmc$async_revdeps/packages, dependencies = NA, recursive = TRUE
```

```

cmc$update()
cmc$async_update()
cmc$check_update()
cmc$async_check_update()

cmc$summary()

cmc$cleanup(force = FALSE)

```

### Arguments

- `primary_path`: Path of the primary, user level cache. Defaults to the user level cache directory of the machine.
- `replica_path`: Path of the replica. Defaults to a temporary directory within the session temporary directory.
- `platforms`: Subset of `c("macos", "windows", "source")`, platforms to get data for.
- `r_version`: R version to create the cache for.
- `bioc`: Whether to include BioConductor packages.
- `cran_mirror`: CRAN mirror to use, this takes precedence over `repos`.
- `repos`: Repositories to use.
- `update_after`: `difftime` object. Automatically update the cache if it gets older than this. Set it to `Inf` to avoid updates. Defaults to seven days.
- `packages`: Packages to query, character vector.
- `dependencies`: Which kind of dependencies to include. Works the same way as the `dependencies` argument of `utils::install.packages()`.
- `recursive`: Whether to include recursive dependencies.
- `force`: Whether to force cleanup without asking the user.

### Details

`cranlike_metadata_cache$new()` creates a new cache object. Creation does not trigger the population of the cache. It is only populated on demand, when queries are executed against it. In your package, you may want to create a cache instance in the `.onLoad()` function of the package, and store it in the package namespace. As this is a cheap operation, the package will still load fast, and then the package code can refer to the common cache object.

`cmc$list()` lists all (or the specified) packages in the cache. It returns a tibble, see the list of columns below.

`cmc$async_list()` is similar, but it is asynchronous, it returns a deferred object.

`cmc$deps()` returns a tibble, with the (potentially recursive) dependencies of packages.

`cmc$async_deps()` is the same, but it is asynchronous, it returns a deferred object.

`cmc$revdeps()` returns a tibble, with the (potentially recursive) reverse dependencies of packages.

`cmc$async_revdeps()` does the same, asynchronously, it returns an deferred object.

`cmc$update()` updates the the metadata (as needed) in the cache, and then returns a tibble with all packages, invisibly.

`cmc$async_update()` is similar, but it is asynchronous.

`cmc$check_update()` checks if the metadata is current, and if it is not, it updates it.

`cmc$async_check_update()` is similar, but it is asynchronous.

`cmc$summary()` lists metadata about the cache, including its location and size.

`cmc$cleanup()` deletes the cache files from the disk, and also from memory.

## Columns

The metadata tibble contains all available versions (i.e. sources and binaries) for all packages. It usually has the following columns, some might be missing on some platforms.

- `package`: Package name.
- `title`: Package title.
- `version`: Package version.
- `depends`: Depends field from DESCRIPTION, or NA\_character\_.
- `suggests`: Suggests field from DESCRIPTION, or NA\_character\_.
- `built`: Built field from DESCRIPTION, if a binary package, or NA\_character\_.
- `imports`: Imports field from DESCRIPTION, or NA\_character\_.
- `archs`: Archs entries from PACKAGES files. Might be missing.
- `reporidir`: The directory of the file, inside the repository.
- `platform`: Possible values: macos, windows, source.
- `needscompilation`: Whether the package needs compilation.
- `type`: bioc or cran currently.
- `target`: The path of the package file inside the repository.
- `mirror`: URL of the CRAN/BioC mirror.
- `sources`: List column with URLs to one or more possible locations of the package file. For source CRAN packages, it contains URLs to the Archive directory as well, in case the package has been archived since the metadata was cached.
- `filesize`: Size of the file, if known, in bytes, or NA\_integer\_.
- `sha256`: The SHA256 hash of the file, if known, or NA\_character\_.
- `deps`: All package dependencies, in a tibble.
- `license`: Package license, might be NA for binary packages.
- `linkingto`: LinkingTo field from DESCRIPTION, or NA\_character\_.
- `enhances`: Enhances field from DESCRIPTION, or NA\_character\_.
- `os_type`: unix or windows for OS specific packages. Usually NA.
- `priority`: "optional", "recommended" or NA. (Base packages are normally not included in the list, so "base" should not appear here.)
- `md5sum`: MD5 sum, if available, may be NA.

- `sysreqs`: For CRAN packages, the `SystemRequirements` field, the required system libraries or other software for the package. For non-CRAN packages it is `NA`.
- `published`: The time the package was published at, in GMT, `POSIXct` class.

The tibble contains some extra columns as well, these are for internal use only.

## Examples

```
dir.create(cache_path <- tempfile())
cmc <- cranlike_metadata_cache$new(cache_path, bioc = FALSE)
cmc$list()
cmc$list("pkgconfig")
cmc$deps("pkgconfig")
cmc$revdeps("pkgconfig", recursive = FALSE)
```

---

`get_cranlike_metadata_cache`

*The R6 object that implements the global metadata cache*

---

## Description

This is used by the `meta_cache_deps()`, `meta_cache_list()`, etc. functions.

## Usage

```
get_cranlike_metadata_cache()
```

## Examples

```
get_cranlike_metadata_cache()
get_cranlike_metadata_cache()$list("cli")
```

---

`meta_cache_deps`

*Query CRAN(like) package data*

---

## Description

It uses CRAN and BioConductor packages, for the current platform and R version, from the default repositories.



## Usage

```
meta_cache_deps(packages, dependencies = NA, recursive = TRUE)

meta_cache_revdeps(packages, dependencies = NA, recursive = TRUE)

meta_cache_update()

meta_cache_list(packages = NULL)

meta_cache_cleanup(force = FALSE)

meta_cache_summary()
```

## Arguments

packages	Packages to query.
dependencies	Dependency types to query. See the dependencies parameter of <code>utils::install.packages()</code> .
recursive	Whether to query recursive dependencies.
force	Whether to force cleanup without asking the user.

## Details

`meta_cache_list()` lists all packages.

`meta_cache_update()` updates all metadata. Note that metadata is automatically updated if it is older than seven days.

`meta_cache_deps()` queries packages dependencies.

`meta_cache_revdeps()` queries reverse package dependencies.

`meta_cache_summary()` lists data about the cache, including its location and size.

`meta_cache_cleanup()` deletes the cache files from the disk.

## Value

A data frame (tibble) of the dependencies. For `meta_cache_deps()` and `meta_cache_revdeps()` it includes the queried packages as well.

## Examples

```
meta_cache_list("pkgdown")
meta_cache_deps("pkgdown", recursive = FALSE)
meta_cache_revdeps("pkgdown", recursive = FALSE)
```

---

 package\_cache

*A simple package cache*


---

## Description

This is an R6 class that implements a concurrency safe package cache.

## Details

By default these fields are included for every package:

- `fullpath` Full package path.
- `path` Package path, within the repository.
- `package` Package name.
- `url` URL it was downloaded from.
- `etag` ETag for the last download, from the given URL.
- `sha256` SHA256 hash of the file.

Additional fields can be added as needed.

For a simple API to a session-wide instance of this class, see [pkg\\_cache\\_summary\(\)](#) and the other functions listed there.

## Usage

```
pc <- package_cache$new(path = NULL)

pc$list()
pc$find(..., .list = NULL)
pc$copy_to(..., .list = NULL)
pc$add(file, path, sha256 = shasum256(file), ..., .list = NULL)
pc$add_url(url, path, ..., .list = NULL, on_progress = NULL,
  http_headers = NULL)
pc$async_add_url(url, path, ..., .list = NULL, on_progress = NULL,
  http_headers = NULL)
pc$copy_or_add(target, urls, path, sha256 = NULL, ..., .list = NULL,
  on_progress = NULL, http_headers = NULL)
pc$async_copy_or_add(target, urls, path, ..., sha256 = NULL, ...,
  .list = NULL, on_progress = NULL, http_headers = NULL)
pc$update_or_add(target, urls, path, ..., .list = NULL,
  on_progress = NULL, http_headers = NULL)
pc$async_update_or_add(target, urls, path, ..., .list = NULL,
  on_progress = NULL, http_headers = NULL)
pc$delete(..., .list = NULL)
```

## Arguments

- `path`: For `package_cache$new()` the location of the cache. For other functions the location of the file inside the cache.
- `...`: Extra attributes to search for. They have to be named.
- `.list`: Extra attributes to search for, they have to in a named list.
- `file`: Path to the file to add.
- `url`: URL attribute. This is used to update the file, if requested.
- `sha256`: SHA256 hash of the file.
- `on_progress`: Callback to create progress bar. Passed to internal function `http_get()`.
- `target`: Path to copy the (first) to hit to.
- `urls`: Character vector or URLs to try to download the file from.
- `http_headers`: HTTP headers to add to all HTTP queries.

## Details

`package_cache$new()` attaches to the cache at `path`. (By default a platform dependent user level cache directory.) If the cache does not exist, it creates it.

`pc$list()` lists all files in the cache, returns a tibble with all the default columns, and potentially extra columns as well.

`pc$find()` list all files that match the specified criteria (`fullpath`, `path`, `package`, etc.). Custom columns can be searched for as well.

`pc$copy_to()` will copy the first matching file from the cache to `target`. It returns the tibble of *all* matching records, invisibly. If no file matches, it returns an empty (zero-row) tibble.

`pc$add()` adds a file to the cache.

`pc$add_url()` downloads a file and adds it to the cache.

`pc$async_add_url()` is the same, but it is asynchronous.

`pc$copy_or_add()` works like `pc$copy_to()`, but if the file is not in the cache, it tries to download it from one of the specified URLs first.

`pc$async_copy_or_add()` is the same, but asynchronous.

`pc$update_or_add()` is like `pc$copy_to_add()`, but if the file is in the cache it tries to update it from the `urls`, using the stored ETag to avoid unnecessary downloads.

`pc$async_update_or_add()` is the same, but it is asynchronous.

`pc$delete()` deletes the file(s) from the cache.

## Examples

```
## Although package_cache usually stores packages, it may store
## arbitrary files, that can be search by metadata
pc <- package_cache$new(path = tempfile())
pc$list()

cat("foo\n", file = f1 <- tempfile())
```

```
cat("bar\n", file = f2 <- tempfile())
pc$add(f1, "/f1")
pc$add(f2, "/f2")
pc$list()
pc$find(path = "/f1")
pc$copy_to(target = f3 <- tempfile(), path = "/f1")
readLines(f3)
```

---

pkg\_cache\_summary      *Functions to query and manipulate the package cache*

---

### Description

pkg\_cache\_summary() returns a short summary of the state of the cache, e.g. the number of files and their total size. It returns a named list.

### Usage

```
pkg_cache_summary(cachepath = NULL)

pkg_cache_list(cachepath = NULL)

pkg_cache_find(cachepath = NULL, ...)

pkg_cache_get_file(cachepath = NULL, target, ...)

pkg_cache_delete_files(cachepath = NULL, ...)

pkg_cache_add_file(cachepath = NULL, file, relpath = dirname(file), ...)
```

### Arguments

cachepath	Path of the cache. By default the cache directory is in R-pkg, within the user's cache directory. See <a href="#">rappdirs::user_cache_dir()</a> .
...	Extra named arguments to select the package file.
target	Path where the selected file is copied.
file	File to add.
relpath	The relative path of the file within the cache.

### See Also

The [package\\_cache](#) R6 class for a more flexible API.

**Examples**

```
pkg_cache_summary()
pkg_cache_list()
pkg_cache_find(package = "forecast")
tmp <- tempfile()
pkg_cache_get_file(target = tmp, package = "forecast", version = "8.10")
pkg_cache_delete_files(package = "forecast")
```

# Index

cranlike\_metadata\_cache, [5](#)

get\_cranlike\_metadata\_cache, [8](#)

meta\_cache\_cleanup (meta\_cache\_deps), [8](#)  
meta\_cache\_deps, [8](#)  
meta\_cache\_deps(), [5](#), [8](#)  
meta\_cache\_list (meta\_cache\_deps), [8](#)  
meta\_cache\_list(), [5](#), [8](#)  
meta\_cache\_revdeps (meta\_cache\_deps), [8](#)  
meta\_cache\_revdeps(), [5](#)  
meta\_cache\_summary (meta\_cache\_deps), [8](#)  
meta\_cache\_update (meta\_cache\_deps), [8](#)  
meta\_cache\_update(), [5](#)

package\_cache, [10](#), [12](#)  
pkg\_cache\_add\_file (pkg\_cache\_summary),  
[12](#)  
pkg\_cache\_delete\_files  
(pkg\_cache\_summary), [12](#)  
pkg\_cache\_find (pkg\_cache\_summary), [12](#)  
pkg\_cache\_get\_file (pkg\_cache\_summary),  
[12](#)  
pkg\_cache\_list (pkg\_cache\_summary), [12](#)  
pkg\_cache\_summary, [12](#)  
pkg\_cache\_summary(), [10](#)  
pkgcache (pkgcache-package), [2](#)  
pkgcache-package, [2](#)

rappdirs::user\_cache\_dir(), [12](#)

utils::install.packages(), [6](#), [9](#)