Package: wand (via r-universe)

September 20, 2024

Type Package

Title Retrieve Magic Attributes from Files and Directories

Version 0.6.0

Date 2019-07-27

Description MIME types are shorthand descriptors for file contents and can be determined from ``magic'' bytes in file headers, file contents or intuited from file extensions. Tools are provided to perform curated ``magic'' tests as well as mapping MIME types from a database of over 1,800 extension mappings.

URL https://gitlab.com/hrbrmstr/wand

BugReports https://gitlab.com/hrbrmstr/wand/issues

Encoding UTF-8

License MIT + file LICENSE

Suggests covr, tinytest

Imports tools

Depends R (>= 3.2.0)

RoxygenNote 6.1.1

Repository https://hrbrmstr.r-universe.dev

RemoteUrl https://gitlab.com/hrbrmstr/wand

RemoteRef HEAD

RemoteSha 1f89bed4a5aba659376ab7f626dc077ee148df39

Contents

get_content_type *Discover MIME type of a file based on contents*

Description

There are a limited number of header "magic" bytes checked directly by this function but cover quite a bit of ground. After that, `guess_content_type()` is called which uses file extension-to-MIME mappings.

Usage

```
get_content_type(path, guess = TRUE, ...)
```

Arguments

`path`	path to a file
`guess`	if `TRUE` (the default), calls `guess_content_type()` if no internal rules match the magic header
`...`	passed on to `guess_content_type()` if guess is `TRUE`

Details

File an issue or PR if more magic-byte-level comparisons are required/desired. If no match is found, ??? is returned (see `guess_content_type()` for how to override this behaviour).

Value

character vector

Examples

```
get_content_type(system.file("extdat", "pass-through", "test.pdf", package="wand"))
```

guess_content_type *Guess MIME type from filename (extension)*

Description

Uses an internal database of over 1,800 file extension-to-MIME mappings to return one or more associated types for a given input path. If no match is found, ??? is returned.

Usage

```
guess_content_type(path, not_found = "???", custom_db = NULL)
```

Arguments

`path`	path to file
`not_found`	MIME type to use when the content cannot be guessed by file type.
`custom_db`	a single data frames each with two columns: `mime_type` and `extension`. These sources will be used along with the built-in sources and will take priority over the built-in sources. Note that the `extensions` should be lower case as they are in the official MIME database.

Value

character vector

Examples

```
guess_content_type(system.file("extdat", "pass-through", "test.pdf", package="wand"))
```

`simplemagic_mime_db` *File extension-to-MIME mapping data frame*

Description

built from the files found in the References

Note

Last updated 2019-07-27.

References

- <https://tools.ietf.org/html/rfc6838> - <https://www.iana.org/assignments/media-types/media-types.xhtml> - <https://s-randomfiles.s3.amazonaws.com/mime/allMimeTypes.json> - <https://cdn.jsdelivr.net/gh/jshttp/mime-db@master/db.json>

`wand` *Retrieve 'Magic' Attributes from Files and Directories*

Description

The 'libmagic' library provides functions to determine 'MIME' type and other metadata from files through their "magic" attributes. This is useful when you do not wish to rely solely on the honesty of a user or the extension on a file name. It also incorporates other metadata from the mime-db database https://github.com/jshttp/mime-db

Some important details

The header checking is minimal (i.e. nowhere near as comprehensive as `libmagic`) but covers quite a bit of ground. If there are content-check types from magic sources that you would like coded into the package, please file an issue and *include the full line(s)* from that linked `magic.tab` that you would like mapped.

Author(s)

Bob Rudis (bob@rud.is)

Index