Package: rmistral (via r-universe)

March 7, 2025

Type Package

Title Experimental R interface to Mistral AI API

Version 0.1.0

Author Sebastian Kranz

Maintainer Sebastian Kranz <sebastian.kranz@uni-ulm.de>

Description Currently implement only OCR capabilties for PDF conversion.

License GPL >= 2.0

Encoding UTF-8

LazyData true

Depends httr, stringi, restorepoint, base64enc

RoxygenNote 7.2.3

Config/pak/sysreqs libicu-dev libssl-dev

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

RemoteUrl https://github.com/skranz/rmistral

RemoteRef main

RemoteSha 352249491c28f9ae74c7fbe9e824ec6d6780d45b

Contents

mistral_api_key 2
mistral_get_file_url 2
mistral_ocr 3
mistral_ocr_save_images 3
mistral_ocr_save_md 4
mistral_upload_file 5
set_mistral_api_key 5

Index **6**

`mistral_api_key` *Retrieve the Mistral API key*

Description

This function retrieves the Mistral API key stored in the R options.

Usage

```
mistral_api_key()
```

Value

The stored API key as a string.

`mistral_get_file_url` *Get a signed URL for file access*

Description

Get a signed URL for file access

Usage

```
mistral_get_file_url(file_id, expiry = 24, api_key = mistral_api_key())
```

Arguments

`file_id`	ID of the file to get URL for
`expiry`	Expiry time in hours, default is 24
`api_key`	Your MistralAI API key

Value

The signed URL information as a list

`mistral_ocr` *Perform OCR using the Mistral API*

Description

This function sends a document (either via URL or file upload) to the Mistral OCR API for text extraction.

Usage

```
mistral_ocr(
  url = NULL,
  file = NULL,
  model = "mistral-ocr-latest",
  include_images = TRUE,
  timeout_sec = 60 * 5,
  api_key = mistral_api_key()
)
```

Arguments

`url` Optional. The URL of the document to be processed.

`file` Optional. A local file path or an uploaded file object.

`model` The OCR model to use. Defaults to ‘"mistral-ocr-latest"‘.

`include_images` Logical. Whether to include images in the response. Defaults to ‘TRUE‘.

`timeout_sec` The timeout for the API request in seconds. Defaults to 300 (5 minutes).

`api_key` Character string. Your Mistral API key. Best set globabally via calling `set_mistral_api_key`.

Value

List containing the result. The field `is_ok` should be TRUE if everything worked nicely. The element `pages` contains the extracted pages.

`mistral_ocr_save_images` *Save extracted images from OCR results*

Description

This function extracts and saves images from OCR results to a specified directory.

Usage

```
mistral_ocr_save_images(ocr, img_dir, overwrite = FALSE)
```

Arguments

`ocr` A list containing OCR results, including extracted images.

`img_dir` Directory where images should be saved.

`overwrite` Logical. If ‘TRUE‘, overwrites existing image files. Defaults to ‘FALSE‘.

Value

The number of images saved.

`mistral_ocr_save_md` *Save OCR results as a Markdown file*

Description

This function saves the extracted OCR text in Markdown format, either as a single file or split by pages.

Usage

```
mistral_ocr_save_md(
  ocr,
  file,
  by_page = FALSE,
  save_images = TRUE,
  overwrite = FALSE,
  img_dir = dirname(file)
)
```

Arguments

`ocr` A list containing OCR results, including extracted text and images.

`file` The file path where the Markdown content should be saved.

`by_page` Logical. If ‘TRUE‘, saves each page as a separate file. Defaults to ‘FALSE‘.

`save_images` Logical. If ‘TRUE‘, also saves extracted images. Defaults to ‘TRUE‘.

`overwrite` Logical. If ‘TRUE‘, overwrites existing files. Defaults to ‘FALSE‘.

`img_dir` Directory where images should be saved. Defaults to the directory of ‘file‘.

mistral_upload_file *Upload a file for OCR processing*

Description

Upload a file for OCR processing

Usage

```
mistral_upload_file(file_path, purpose = "ocr", api_key = mistral_api_key())
```

Arguments

`file_path`	Path to the local file to upload
`purpose`	Purpose of the upload, default is "ocr"
`api_key`	Your MistralAI API key

Value

The file metadata as a list

set_mistral_api_key *Set the Mistral API key*

Description

This function sets the Mistral API key, either from a provided string or by reading from a file.

Usage

```
set_mistral_api_key(key = NULL, file = NULL)
```

Arguments

`key`	Optional. The API key as a string.
`file`	Optional. A file path containing the API key.

Index

mistral_api_key, 2
mistral_get_file_url, 2
mistral_ocr, 3
mistral_ocr_save_images, 3
mistral_ocr_save_md, 4
mistral_upload_file, 5

set_mistral_api_key, 5