Package 'tesseract'

November 20, 2023

Type Package

Title Open Source OCR Engine

Version 5.2.1

Description Bindings to 'Tesseract':
a powerful optical character recognition (OCR) engine that supports over 100 languages.
The engine is highly configurable in order to tune the detection algorithms and
obtain the best possible results.

License Apache License 2.0

URL https://docs.ropensci.org/tesseract/ (website)

https://github.com/ropensci/tesseract (devel)

BugReports https://github.com/ropensci/tesseract/issues

SystemRequirements Tesseract >= 3.03 (libtesseract-dev /
tesseract-devel) and Leptonica (libleptonica-dev /
leptonica-devel). On Debian you need to install the English
training data separately (tesseract-ocr-eng)

Imports Rcpp (>= 0.12.12), pdftools (>= 1.5), curl, rappdirs, digest

LinkingTo Rcpp

RoxygenNote 7.2.3

Suggests magick (>= 1.7), spelling, knitr, tibble, rmarkdown

Encoding UTF-8

VignetteBuilder knitr

Language en-US

NeedsCompilation yes

Author Jeroen Ooms [aut, cre] (<https://orcid.org/0000-0002-4035-0289>)

Maintainer Jeroen Ooms <jeroen@berkeley.edu>

Repository CRAN

Date/Publication 2023-11-20 15:10:03 UTC