

Documentación multilingüe en paquetes de R.

Elio Campitelli

Palabras clave: traducción - multilingue - documentación

¿Quieres leer la documentación de base: `:mean()` (parcialmente) en español? Instala [rhelpi18n](#) y el módulo de traducción de ejemplo base.es:

```
# install.packages("pak")
pak::pak(c("eliocamp/rhelpi18n", "eliocamp/base.es"))
```

Y luego ejecuta este código:

```
library(rhelpi18n)
Sys.setenv(LANGUAGE = "es")
help(mean)
```

```
## Media Aritmética
##
## Description:
##
##     Función genérica para la media aritmética (recortada)
##
## Usage:
##
##     mean(x, ...)
##
##     ## Default S3 method:
##     mean(x, trim = 0, na.rm = FALSE, ...)
##
## Arguments:
##
##     x: an R object.  Currently there are methods for numeric/logical
##        vectors and date, date-time and time interval objects.
##        Complex vectors are allowed for 'trim = 0', only.
##
##     trim: la fracción (0 a 0.5) de observaciones para ser recordatas
##           del final de \codex antes de computar la media. Valores de
##           trim pro fuera de ese rango se consideran al punto más
##           cercano.
##
##     na.rm: un valor lógico que evalua a 'TRUE' o 'FALSE' indicando si
##            valores 'NA' deben ser removidos antes de realizar el
##            cómputo.
##
##     ...: further arguments passed to or from other methods.
##
... (27 more lines)
```

¡Hermoso!

¿Por qué harías esto?

Aunque el inglés es la lengua internacional de facto, existen [ventajas](#) de incluir recursos en diferentes idiomas. Al reducir las barreras lingüísticas, acabamos teniendo una comunidad más grande, fuerte y creativa, con más ideas y recursos. Esta hegemonía del inglés se refleja en R por el uso del inglés en su documentación, como las páginas de ayuda, los nombres de las funciones y los nombres de los argumentos (¿por qué usar `mean()` en lugar de `Mittelwert()`?).

La documentación de los paquetes *puede* escribirse en cualquier idioma, y CRAN admite documentación en otros idiomas utilizando el campo “Language” (Idioma). Pero aún así, la gran mayoría de los paquetes están documentados en inglés y hay [un pequeño número de paquetes documentados en otros idiomas](#) aparentemente para llegar más fácilmente a su público objetivo. Por ejemplo, el paquete [labstatR](#) sirve de complemento al libro italiano “Laboratorio Di Statistica Con R” y está parcialmente documentado en italiano (utiliza nombres de funciones y argumentos en inglés). Del mismo modo, [chilemapas](#) proporciona mapas simplificados para Chile, con documentación completa en español, incluidos los nombres de las funciones. [rOpenSci censo2017](#) por [Mauricio Vargas](#) también está totalmente documentado en español.

Aunque estos paquetes son más accesibles para su público principal, son mucho menos accesibles para la comunidad en general. Quienes que no dominan el idioma pueden tener dificultades para descubrir y utilizar las funciones que ofrecen estos paquetes. Quienes desarrollan estos paquetes se enfrentan al dilema de hacer que su paquete sea inaccesible para su público objetivo o aislarlo del resto del ecosistema de R.

Próximos pasos

Este paquete es un proyecto del nuevo [Grupo de Trabajo de Documentación Multilingüe del R Consortium](#). Este Grupo de Trabajo surgió luego de varias conversaciones durante el [R Project Sprint 2023](#) y supervisará la implementación del soporte de documentación multilingüe en R y organizará los esfuerzos de traducción de la comunidad.

El proyecto está en sus primeras fases por lo que las cosas cambian rápido. Aún estamos experimentando con distintas formas de almacenar las traducciones y de aplicar sustituciones de texto. También tiene varias limitaciones, como que aún no traduce los encabezados de sección. También queremos añadir una indicación clara de que el texto es una traducción y una forma conveniente de acceder rápidamente a la versión original. Parte del flujo de trabajo de las traducciones suele consistir en traducciones automáticas que posteriormente revisa una persona. Por ello, estamos considerando la posibilidad de añadir metadatos a las cadenas para señalar los casos de traducciones automáticas que no han sido revisadas por una persona o las traducciones que podrían estar ligeramente desactualizadas (¿sería una pena no mostrar una traducción porque alguien arregló un pequeño error de tipeo en el texto original y aún no había sido revisada!)

Si deseas probar el paquete e informar de cualquier problema, o si tienes ideas sobre cómo mejorarlo, puedes participar en el debate en los [issues abiertos](#) del paquete y uniéndote al [Grupo de trabajo de documentación multilingüe](#). ¡Ayúdanos!