

Reemplazando la torre de Babel: herramientas para documentos multilingües en R

Maëlle Salmon, Yanina Bellini Saibene, Paola Corrales, Elio Campitelli

Abstract

La comunidad de usuarios y usuarias de R es global y por lo tanto multilingüe. Quienes producimos contenidos sobre R, ya sean sitios de documentación, libros, entradas de blog, adaptamos frecuentemente nuestros documentos a audiencias en varios idiomas. En rOpenSci tenemos un [proyecto fructífero de publicación multilingüe](#), para el cual traducimos nuestra [guía de desarrollo de paquetes](#) y otros materiales de inglés a español. También se está realizando esta traducción al portugués. Llevar a cabo este trabajo necesitó desarrollar no solo procesos humanos, sino también dos paquetes de R: babelquarto para construir sitios y libros Quarto multilingües, y babeldown para utilizar la API de traducción automática DeepL para acelerar el proceso de traducción de documentos Markdown. Babelquarto y babeldown resultaron útiles más allá de nuestro proyecto. En esta charla presentaremos los dos paquetes, sus funcionalidades y límites.

babelquarto: para construir sitios y libros Quarto multilingües

Nuestra guía de desarrollo está publicada con Quarto, porque es una herramienta moderna que al contrario de, por ejemplo, Hugo, es bastante fácil de aprender y además útil para otras aplicaciones como reportes y presentaciones. Desafortunadamente, en Quarto nativo no se pueden construir sitios y libros multilingües: sitios y libros donde cada página tiene un enlace a su versión en otros idiomas.

Por eso decidimos crear un paquete que permite construir tales libros y sitios con Quarto. Basándose en documentos con distintas extensiones dependiendo del idioma y en campos de configuración específicos, babelquarto construye las diferentes versiones de un libro o sitio, y añade los enlaces hacia otros idiomas.

El paquete funciona perfectamente para nuestra guía de desarrollo de paquetes. También es utilizado por otras comunidades y proyectos, resultando en contribuciones que hacen al paquete aún más colaborativo y conveniente para otros usos.

babeldown: para traducir documentos Markdown con DeepL

Necesitamos reducir los esfuerzos humanos con la automatización. Primero, en vez de traducir documentos desde cero, nos apoyamos en una traducción automática hecha por

REEMPLAZANDO LA TORRE DE BABEL: HERRAMIENTAS PARA DOCUMENTOS MULTILINGÜE EN R

DeepL, lo que nos permite enfocar nuestro trabajo de traducción manual en aspectos como elegir las palabras más adecuadas para nuestra comunidad o usar formulaciones neutrales en cuanto al género. Segundo, para poder automatizar el proceso al máximo, no queremos copiar y pegar desde una herramienta externa. Tercero, nuestros documentos están preparados en Markdown por lo que la traducción automática debe respetar esta sintaxis. Con esos tres criterios creamos el paquete de R, babeldown, que envía el contenido de documentos Markdown a la API de traducción DeepL y escribe el resultado en documentos Markdown correctos. Además, con babeldown se pueden actualizar traducciones traduciendo únicamente las partes que cambiaron, ahorrando tiempo para las personas que traducen y dinero si se paga el uso de la API.

Lecciones aprendidas

La infraestructura creada se integra adecuadamente a flujos de trabajo con R, Git y GitHub y ha sido utilizada exitosamente por rOpenSci para traducir contenido en Inglés, Portugues, Español y Francés. También está siendo utilizada por otros proyectos como R4Epi.

Recursos

- <https://docs.ropensci.org/babelquarto/>
- <https://docs.ropensci.org/babeldown/>
- <https://masalmon.eu/talks/2023-11-17-bcn/>
- <https://ropensci.org/commcalls/nov2023-multilenguaje/>