

# Mejorando los llamados a paquetes

**Palabras clave:** annotater, librerías, SQL

## Resumen

Los lenguajes de programación como R se benefician de la gran variedad de paquetes que extienden sus capacidades, y trabajar con código nos permite desarrollar flujos de trabajo transparentes y reproducibles. Sin embargo, la enorme diversidad de paquetes disponibles puede ser abrumadora, al grado que se dificulta saber: para qué sirven los diferentes paquetes, de dónde se obtuvieron, o por qué se están utilizando en algún proyecto. El repositorio principal de paquetes para R (CRAN, The Comprehensive R Archive Network) hoy en día contiene más de 19,000 librerías, y en otros repositorios existen muchas más.

Al trabajar con *scripts* y cargar librerías para nuestros proyectos, es posible hacer anotaciones y comentarios sobre los paquetes que se incluyen en un archivo .R, .Rmd o .Qmd. Para agregar contexto a un archivo que está llamando paquetes con la función `library()`, el paquete {annotater} aporta herramientas y funciones que pueden insertar información como la fuente, versión, título, o los datos o funciones que se están utilizando de cada paquete. Por ejemplo: la línea `library(purrr)` en un script se puede convertir en `library(purrr) # CRAN v1.0.1`. Para suplementar las posibles anotaciones que se pueden añadir a los llamados, este trabajo explora lo que otras personas están comentando sobre los paquetes que utilizan.

Se utilizó la plataforma BigQuery de Google, que captura semanalmente todo el contenido disponible en repositorios públicos de GitHub, uno de los principales portales para alojar código. De aproximadamente medio millón de archivos con código de R, casi 4000 incluían al menos una llamada a la función `library()` seguida por un comentario (indicado por el símbolo de comentarios “#”). Los comentarios sobre los paquetes en general se pueden asignar en tres usos distintos: notas sobre la instalación del mismo (mencionando instrucciones sobre repositorio o versión), para qué se está cargando el paquete, y cuáles funciones se van a usar en pasos subsiguientes. Fue interesante encontrar un gran número de comentarios haciendo aclaraciones sobre paquetes que conforman el universo ‘tidyverse’. Se detectaron comentarios en al menos ocho idiomas distintos, aunque el inglés fue el lenguaje más común.

Esta exploración inicial de comentarios sustenta la idea original detrás del paquete {annotater}; no siempre es evidente para qué se requieren diferentes paquetes. Además, dejar registro de versión y fuente puede servir para mejorar la reproducibilidad de un script sin necesidad de otras herramientas más elaboradas.