**rtweet: Collecting and analyzing Twitter data**

Michael W. Kearney

1 School of Journalism, Informatics Institute, University of Missouri

**Statement of need**

Following the announced (2016) deprecation of the *twitteR* package (Gentry, 2013), R users seeking to interact with Twitter APIs have been encouraged to use the *rtweet* package. Use of the up-to-date and actively-maintained *rtweet* package is especially important in light of changes to Twitter’s APIs since 2016. Most notably is the increased character limit (from 140 to 280) for Twitter statuses (Rosen & Ihara, 2017). In addition to providing an updated interface with similar functionality to *twitteR*, allowing R users to communicate with various endpoints from Twitter’s REST API, the *rtweet* package also provides support for communicating with Twitter’s stream API.

**Summary**

Interest in Twitter data continues to grow, but for many the task of actually collecting and analyzing data via Twitter APIs remains daunting. For example, in order to interact with Twitter’s APIs users must, in addition to identifying and digesting the relevant information from Twitter’s developer documentation, build/send/receive requests, manage rate limits, and wrangle nested and real-time response objects into analysis-friendly data structures. Fortunately, the *rtweet* R package (Kearney, 2018a) is designed to simplify these processes, making interacting with Twitter’s APIs more accessible to a wider range of users.

The main goals of the *rtweet* package are two-fold. The first goal is to make interacting with Twitter’s APIs more approachable and streamlined for less computationally-experienced users. The second goal is to assist in the analysis of Twitter data via converting information returned by Twitter’s APIs into tabular data structures and providing several convenience functions for common analytical techniques such as examining Twitter networks or the frequency of tweets over time. In short, although it is certainly possible for users to write their own Twitter API wrapper functions, the heavy-lifting done by *rtweet* to (a) streamline the building, authorizing, and sending of API requests, (b) wrangle deeply nested JSON data into tabular structures, and (c) provide convenience functions for for relevant and popular analytical techniques, make it a valuable contribution in the area of collecting and analyzing Twitter data.

Although *rtweet* provides some coverage to user-context-behaviors (e.g., posting statuses, liking tweets, following users, etc.), the primary audience for the package to date has been researchers. Accordingly, *rtweet* has been featured in numerous popular press (e.g., Bajak & Wu, 2019; Machlis, 2019; Riley, 2019) and academic publications (e.g., Bossetta, 2018; Bradley & James, 2019; Buscema, Ferilli, Massini, & Zavarrone, 2018; Erlandsen, 2018; Gitto & Mancuso, 2019; Kearney, 2018b, 2019; Li, Chamrajnagar, Fong, Rizik, & Fu, 2019; Lutkenhaus et al., 2019a, 2019b; Molyneux, Lewis, & Holton, 2018; Tsoi et al., 2018; Unsihuay, 2018; Valls et al., 2017; Wu & Ying, 2018).
References


