

UConn

CENTER FOR MHEALTH
AND SOCIAL MEDIA

Using Rtweet and NCapture to Gather Twitter Data

Molly E. Waring, PhD

Associate Professor, Department of Allied Health Sciences
Director of Methodology Core, UConn Center for mHealth and Social Media

Email molly.waring@uconn.edu

Twitter [@DrMollyWaring](https://twitter.com/DrMollyWaring)

Using social media in health research

- Recruit for trials or observational studies
- Deliver behavioral interventions
- Evaluate social media-delivered interventions
- Observational studies of social media content



<https://mhealth.inchip.uconn.edu/events/>

"An Overview of Methods for Capturing and Examining Social Media Data for Health Research" <https://www.youtube.com/watch?v=B-wMT151NrY>

Today: Collecting Twitter Data

- Considerations
 - Methodological
 - Ethical
 - Practical
- Walk through how to use two tools to collect Twitter data
 - NCapture
 - Rtweet

Methodological Considerations

- Data collection (and analysis) should stem from research question
- Interdisciplinary team should include subject matter expertise & data science/analytics
 - Knowledge of search terms, hashtags, & jargon/slang used by target population
 - Lots of data management required to turn raw data into information
- Sampling: prospective vs retrospective; seasonality; day of week; popular events
- Interpretation and generalizability:
 - What proportion of your population of interest uses Twitter?
 - How does your target population use Twitter?

Ethical Considerations

- We have an ethical obligation to protect the people who contribute research data regardless of whether the project meets criteria for human subjects research
- IRB approval required for intervention study
 - Include details in consent form and discussion about what data will be collected (and what data won't be collected)
 - Sample consent forms available through Connected and Open Research Ethics (CORE) <https://thecore.ucsd.edu/>
- Tweets posted publicly vs privately; expectations of privacy?
- How to handle suicidal ideation or other participant safety concerns or reportable disclosures?

<https://pubmed.ncbi.nlm.nih.gov/29671135/>; <https://pubmed.ncbi.nlm.nih.gov/29942634/>;
<https://pubmed.ncbi.nlm.nih.gov/30698793/>; <https://pubmed.ncbi.nlm.nih.gov/31855185/>

Practical Considerations

- Which tool best suits your data needs?
 - Does someone on your team know R? Does your team have access to NVivo?
- Authorize through your Twitter account - so be careful about not capturing tweets from private accounts who have accepted you as a follower
- Twitter's API will change - be prepared to adapt

NCapture

- Free browser extension for Chrome and Internet Explorer to NVivo
- <https://help-nv.qsrinternational.com/12/win/v12.1.96-d3ea61/Content/ncapture/ncapture.htm>
- <https://help-nv.qsrinternational.com/12/win/v12.1.90-d3ea61/Content/ncapture/social-media-data.htm>
- Twitter, Facebook, YouTube, webpages
- Import data into NVivo for analysis or to export to other statistical software
- Note, not always consistent # tweets downloaded or how far back tweets are captured

rtweet



- R package
- Download/install R <https://www.r-project.org/>
- <https://cran.r-project.org/web/packages/rtweet/vignettes/intro.html>
- <https://cran.r-project.org/web/packages/rtweet/index.html>
- Search by keywords (hashtags), timelines (tweets), liked tweets, and more
- Note, keyword/hashtag searches limited to past 7 days